7-21-2023

Phylogeny and Taxonomy of Genus Physaria in North America

Binoshi Hettihewa
bjhg4n@umsystem.edu

PHYLOGENY AND TAXONOMY OF GENUS *PHYSARIA* IN NORTH AMERICA

by

Binoshi Hettihewa

Bachelor of Science Honors in Plant Science, University of Colombo Sri Lanka, 2020

A Thesis Submitted to The Graduate School of the University of Missouri-St. Louis
in partial fulfillment of the requirements for the degree

Master of Science in Biology
With an emphasis in Ecology, Evolution, and Systematics

August 2023

Advisory Committee

Nathan Muchhala, Ph.D.
Chairperson

Christine Edwards, Ph.D.

Aimee Dunlap, PhD.

ACKNOWLEDGMENTS

I'm deeply indebted to my supervisor Dr. Christine Edwards for her invaluable guidance, advice, feedback, patience, and immense support given throughout the time. Her understanding and constant appreciation has motivated me to become a better student. I'm extremely grateful for my second supervisor Dr. Nathan Muchhala for his invaluable support, advice, and encouragement and discussions in his lab helped me to deepen my knowledge. I'm extremely grateful to Dr. Aimee Dunlap for her support and guidance throughout my research. I would like to express my deepest gratitude to Brock Mashburn for teaching me all the new things in the lab and his generous support, encouragement, patience, and friendship given throughout the period. I extend my sincere thanks to Burgund Bassuner for training me in the molecular lab and her enthusiasm given throughout the period and making the lab a very welcoming place. I'm also grateful for Alexander Linan for his support and making time to discuss ideas and feedback. I would like to extend my sincere thanks to Prof. Steve O' Kane for his feedback and guidance. I would also like to thank all the faculty members at Biology Department at University of Missouri St. Louis for their support and advice during the period and motivating me to become a better scientist.

I'm immensely grateful for the Harris World Ecology Center of University of Missouri St. Louis for funding and making this project a reality. I would also like to thank all other lab members in my lab, Eduardo Aguirre-Mazzi, Brigette Williams, Rachel Brant and Juan Carlos Penagos for being supportive and friendly.

It was such a privilege to work with dynamic and international environment at Missouri Botanical Garden. I'm grateful for the staff of the Missouri Botanical Garden, especially the staff at the library for being helpful whenever I need.

I had the pleasure of working with fellow graduate students at the Department of Biology, Belen Alvestegui, Prasai Raj, Sebastian Forward, Gorge Todd, Emily Beahm, Jordan Hathway, Andrea Trigueros, Kristen Rosamend, Jhon Bender, Ketra Oketcho and Daniel Ocana. Many thanks to Pavithra Madamarandawala, Yughandara Eriyagama and Shanaka De Silva for their kindness and support throughout the time. Finally, a very affectionate thanks to my parents and sister for their kindness, courage, love and support for me whenever possible.

ABSTRACT

Physaria is a genus of ~108 species belonging to family Brassicaceae that is predominantly distributed in Western North America, but one species occurs in Arctic Russia and Northern Canada and several species occur in South America. Regardless of the vast number of species in the genus, the genus lacks a well-resolved phylogeny representing many taxa, partially because phylogeny reconstruction is complicated by the fact that many species of *Physaria* vary in chromosome numbers and ploidy levels. In chapter 1, we review how polyploids are formed and become established and summarize what is known about variation in chromosome number and ploidy in Brassicaceae and in the genus *Physaria*. In Chapter 2, we extracted DNA representing 84 species of *Physaria* species and employed a 2b-RAD sequencing technique to generate data for phylogeny reconstruction. The specific goals of the study were 1) to reconstruct the phylogeny of *Physaria* and assess whether species relationships proposed in early monographs by Payson (1921) and Rollins and Shaw (1973) based on morphology correspond to the current species relationships revealed through the molecular phylogeny representing 86 species of *Physaria*, 2) to investigate the monophyly of species represented by multiple accessions in the resulting phylogeny, and 3) to investigate how the inclusion of polyploid taxa affects the topology of the phylogeny, which may help shed light on the origins of polyploid taxa. The resulting phylogeny had species from Mexico and Texas at the base of the tree, suggesting that the genus originated in southern North America, although additional outgroups and a formal biogeographic analysis are needed to confirm this result. The species in the phylogeny were grouped into two main clades, one containing species

predominantly from eastern North America, and one containing species predominantly from western North America. Except for the group 1 species proposed by Rollins and Shaw in 1973, none of proposed groups of species in monographs formed clades. Instead, the resulting phylogeny grouped species collected from nearby locations irrespective of their taxonomic placement, suggesting a strong biogeographical affinity towards species groupings, possibly due to hybridization within geographic locations. Only a handful of species were monophyletic; in the eastern clade, *P. recurvata*, *P. gracilis*, *P. angustifolia*, *P. globosa*, and a new species, *P. ouachitensis* formed monophyletic species in phylogenies in which the polyploid species were present and absent. In the western clade, *P. brassicoides*, *P. pruinosa*, *P. valida*, *P. parvula*, *P. pulvinata* and *P. intermedia* formed monophyletic species when the polyploids were removed. The topology varied depending on whether polyploids were included, suggesting that some species may be hybrids or allopolyploids. Conducting additional chromosome counts, identifying hybrid and allopolyploid taxa, and reconstructing the evolution of ploidy levels is an important area for future studies to understand how they have affected diversification in the group. Overall, the current study resulted a well resolved phylogeny with many taxa of *Physaria*, which is useful for future studies on understanding the evolutionary history, character evolution, and biogeography of the genus.

Key words: *Physaria*, 2b-RAD, polyploidy, molecular phylogenetic analysis, species monophyly

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

CHAPTER 1 MECHANISMS, ESTABLISHMENT, AND ECOLOGICAL AND EVOLUTIONARY CONSEQUENCES OF POLYPLOIDY IN PLANTS

Abstract

Polyploidization, which is when an organism has more than two complete sets of chromosomes, has been identified as a major mechanism of species diversification. All extant plants have undergone at least one whole genome duplication event in the past. The main mechanisms of polyploid formation are via unreduced gamete formation, somatic doubling, and rarely, by polyspermy Autopolyploidy and allopolyploidy are the two types of polyploid formation, which differ in whether the polyploid arose from one species or from hybridization between two species. An allopolyploid plant is one that arose through interspecific hybridization and experienced either genome doubling within a single individual or self-fertilization involving unreduced gametes. An autopolyploid is derived from a spontaneous increase in the number of chromosomes sets in a single species. Newly formed polyploids face challenges when they try to establish in a mixed-cytotype population due to obstacles such as meiotic irregularities, reduced fertility, altered gene dosage, instantly altered physiological properties, and minority cytotype disadvantage, Polyploids overcome minority cytotype exclusion via different mechanisms, such as recurrent formation of polyploids, assortative mating and asynchronous flowering. We also focus on polyploidy in the Family Brassicaceae, which has experienced several polyploidization events in its origin and throughout its evolutionary history that have affected its diversification. We also examine patterns of polyploidy in the genus *Physaria*, which is a polyploid-rich genus that offers many opportunities to investigate the origins, establishment, and evolutionary consequences of polyploidy.

Key words: Autopolyploidy, allopolyploidy, minority cytotype exclusion, speciation, triploid bridge, triploid block, Brassicaceae

1.1 Introduction

Polyploidy is the condition where an organism has more than the two sets of nuclear chromosomes found in diploid organisms (Husband et al., 2013). Polyploids are denoted based on the base chromosome number of the species. The sporophytic chromosome number is denoted as 2n irrespective of the ploidy level, as they can be diploid, triploid, tetraploid or higher ploidy (Heslop_Harrison et al., 2023) and the gametophytic chromosome number is denoted by n. The base chromosome number is denoted by x and ploidy levels can be denoted by 2x, 3x, 4x or higher for diploids, triploids, and tetraploids, respectively. Species may have a series of polyploids based on a base chromosome number. For example, *Rumex* has a series with base chromosome number of x=10, which runs from 2n=2x=20, through 2n=4x=40 and up to 2n=20x=200. Another example of polyploid series with a base chromosome number is genus *Chrysanthemum* (x=9) with 2n=18, 36, 54, 72, and 90 (Ramsey and Schmeske 1998; Heslop_Harrison et al., 2023). The main mechanisms of polyploid formation are via unreduced gamete formation, somatic doubling, and rarely, by polyspermy. Generally, scientists recognize two main types of polyploids: autopolyploids form through unreduced gametes or chromosome doubling within a single species, whereas allopolyploids are formed through interspecific hybridization, followed by an increase in chromosome number via unreduced gametes or chromosome doubling.

Polyploids are found in a broad range of taxonomic groups, such as insects, fishes, amphibians, and reptiles, but polyploidy is much more common in plants (Van de peer et al., 2017). All angiosperms are thought to be derived from a polyploidization event at the base of the

angiosperm tree of life (Wood et al., 2009), and many subsequent polyploidization events have occurred throughout the evolutionary history of many plant clades (Wood et al., 2009). Given the high frequency of polyploidy and its importance in the evolutionary history of plants, the goal of this review is to outline the history of research on polyploidy in plants, the ways in which polyploids form, establish, and speciate, and to review the ecological and evolutionary consequences of polyploidy both in Angiosperms and specifically in the Brassicaceae family. The specific aims of this review are to describe 1) early work on polyploids, 2) what is known about the formation of polyploids, 3) how to differentiate between autopolyploids and allopolyploids, 4) the factors affecting the establishment of neo polyploids in a mixed ploidy population, 5) the ecological and evolutionary consequences associated with polyploidization, and 6) the mechanisms of reproductive isolation of polyploids in a mixed ploidy population. We also focus on polyploidy in the plant family Brassicaceae and its effect on the subsequent diversification of the family.

1.2 Early work on polyploidy

Polyploidy was first discovered in 1907, and the term polyploidy was first coined by Straburger in 1910. Early work included research by Winkler (1916), who described the experimental production of polyploids in Tomato, *Lycopersicon esculentum*, and Nightshade, *Solanum nigrum*. Other early work on polyploidy was conducted by Hugo de Vries and G. Ledyard Stebbins (Yves Van de peer et al., 2017, Husband et al., 2013). De Vries's work on evening primrose, *Oenothera lamarkiana* (Oenothera), caught special attention because it sought to understand why some plants were generally larger than their parent plant, which was the called "gigas effect". He found that the gigas mutant had twice the somatic chromosome number of wild type plants and argued that the increased chromosome number was the reason behind the

morphological differences between the two morphs. Subsequent work on artificial polyploids has since supported the conclusion that the gigas effect is a result of chromosome doubling (Briggs and Walters, 1997).

In 1917, Winge discovered the existence two types of polyploids, allopolyploids and autopolyploids, and revealed that allopolyploids are formed through hybridization and subsequent chromosome doubling, whereas autopolyploids are formed through chromosome doubling within single species. The terms autopolyploidy and allopolyploidy were later coined by Kihara and Ono (1926). Many studies have since supported the role of hybridization in the formation of allopolyploids; for example, diploid *Nicotinana glutinosa* (2n=25) was crossed with a tetraploid *N. tabacum* (2n=48) and produced a fertile hexaploid (2n = 72), and hybridization between *Brassica oleracea* (Cabbage) 2n=18 and *Raphanus sativus* (radish) 2n=72, giving rise to a fertile tetraploid *Raphanobrassica* (2n = 36) (Brigg and Walters, 1997).

1.3 The formation of polyploids

Polyploids can arise via several different mechanisms. The two most common ways that polyploids arise are through somatic doubling and unreduced gametes. Somatic doubling is where a plant experiences a spontaneous doubling of chromosome number, which can occur in zygotic, embryonic or sporophytic tissue. This produces individuals that have cells with different ploidy levels. Somatic doubling can produce viable polyploid offspring if gamete formation occurs in the parts of the plant that are duplicated (Ramsey and Schemske, 1998). For example, hybridization between two species can leave the resulting offspring infertile because of problems with chromosomal pairing at meiosis, but subsequent somatic doubling can restore fertility

because each set of chromosomes has another homologous set with which to pair (Ramsey and Schemske, 1998).

The other main mechanism of polyploid formation is through unreduced gametes (eggs and pollen), which occurs during micro- and megasporogenesis. In comparison to typical meiosis that produces four haploid cells (reduced gametes, "n"), an unreduced gamete is formed when meiosis results in two cells that have the somatic chromosome number of the parent ("2n") (Oleszczuk et al., 2019). Unreduced gamete formation occurs across the plant tree of life and occurs in about 0.1–2% of the gametes produced (Ramsey and Schemske, 1998). Unreduced gametes can produce polyploid offspring either because they fuse with other unreduced gametes, producing a tetraploid offspring, or when they fuse with reduced gametes, in which case they can sometimes form a triploid (i.e., a "triploid bridge"; Schinkel et al., 2017). These triploids can be partially fertile, producing n, 2n and 3n gametes, resulting in diploid, triploid, or tetraploid seeds by backcrossing to their parental diploids (Bretagnolle, 2001, Schinkel et al., 2017). However, the fusion of unreduced gametes with reduced gametes can be prevented by a triploid block, where one pollen nucleus will fuse with the egg cell and form a triploid zygote but the other sperm cell fails to fertilize the vegetative nuclei/polar nuclei to form the endosperm. This will block triploid seed formation, as they contain a triploid embryo with an unbalanced ratio of maternal to paternal chromosomes in the endosperm. Overall, the frequency of unreduced gamete formation is an important factor affecting the rate of polyploid formation and the types of polyploids being formed (Ramsey and Schemske, 1998)

Unreduced gamete production is primarily due to the failure of meiotic processes, which can occur because of many different types of failures and at different points is meiosis. Meiotic failures that produce unreduced gametes include pre-meiotic doubling, premature cytokinesis, a

mutation in synapsis, a lack of spindle fiber segregation, post meiotic fusion and defects in the mitotic cell cycle (Bretagnolle 2001, Sora *et al*., 2016). Meiosis that results in unreduced gametes is also called meiotic restitution because it produces the same number of chromosomes in the gamete as is found in its parent (Ramanna and Jacobsen 2003; Oleszczuk *et al*., 2019). Generally, meiotic restitution can occur at two different stages during meiosis. First division restitution (FDR) occurs when meiosis is unable to separate homologous chromosomes, whereas second division restitution (SDR) occurs when meiosis is unable to separate sister chromatids (Bretagnolle and Thompson 1995; Oleszczuk *et al*. 2019).

Both environmental and genetic factors may increase the failure rate of meiotic processes that result in unreduced gamete formation. Adverse environmental conditions that affect meiosis can accelerate unreduced gamete production; for example, stress related to adverse temperatures, photoperiods and water availability are known to be drivers of unreduced gamete formation (Schinkel et al., 2017). Hybridization is one main genetic factor leading to unreduced gamete formation; hybrids often have non-compatible chromosomes that experience irregular pairing during meiosis, leading to chromosome segregation failure. Triploids often produce increased numbers of unreduced gametes (Sora et al., 2016).

Several techniques are available to identify unreduced gametes. Traditionally, unreduced pollen grains have been identified by their size because their diameter is typically 30-40% larger than reduced pollen (Ramsey and Schmeske, 1998). These differences can be detected using volume-based particle size analysis techniques. Another approach is to analyze the microspores from post-meiotic pollen mother cells to detect dyads, which indicates the failure of the meiosis in pollen mother cells (Sora et al., 2016).

Several studies have investigated the role of unreduced gamete production in producing polyploids in natural populations of plants. Bretagnolle (2001) studied the frequency of unreduced pollen production and polyploid seed production in a diploid population of *Anthoxanthum alpinum* by observing the frequency of large (i.e., unreduced) pollen and screening the ploidy of the seeds produced. They found evidence of unreduced gametes, but no tetraploids were produced. Instead, the found several triploid seeds, suggesting that the species can overcome the triploid block. Schinkel et al (2017) aimed to understand the formation of polyploids in natural populations of the alpine plant *Ranunculus kuepferi*. The ploidy level of seeds was analyzed to identify the endosperm and embryo ploidy level, and apomictic seeds were differentiated from sexual seeds. Different ploidy types were identified by studying the histograms of the ratio of embryo/endosperm genome size using flow cytometry. They analyzed whether seeds resulted from a female triploid bridge (unreduced egg cell fertilized by reduced pollen), male triploidization (reduced egg fertilized by unreduced pollen), polyhaploids (reduced egg cell of tetraploid without fertilization), disturbed sexuals (irregular male or female meiosis resulting in aneuploidy), biparental polyploidization (unreduced egg cells fertilized by unreduced pollen). Results showed that the formation of polyploids in *Ranunculus kuepferi* plants in the wild occurred via unreduced eggs (female triploid bridge) and biparental polyploidization events in *R. kuepferi.* Tetraploids of *R. kuepferi* were found to reproduce mainly via apomixis (Schienkel et al., 2017).

Finally, the third mechanism that can produce polyploids is polyspermy, which is the fertilization of the ovule by more than one sperm, which occurs rarely. Polyspermy is naturally prevented by the mechanisms that ensure double fertilization of angiosperms. However, sometimes two pollen nuclei can fuse with one egg cell and form triploids, which is called hetero

fertilization. If the pollen is from the same male plant, it produces a biparental triploid, whereas if the pollen comes from two different parents, it produces a triparental triploid plant. However, polyspermy occurs very infrequently because as soon as the pollen nuclei fertilize the egg cell, the cell wall forms, acting as a slow polyspermy block that acts on both the egg cell and the central cell to ensure the double fertilization (Toda and Okamoto, 2020).

1.3.1 Autopolyploidy and allopolyploidy

Allopolyploidy or autopolyploidy can be defined with respect to the mode of origin, or by cytological or genetic criteria. According to the mode of origin, autopolyploids arise within single populations or through hybridization between cytotypes of a single species. Allopolyploids arise due to interspecific hybridization, followed either by genome doubling or through the production and fusion of unreduced gametes. According to cytological criteria, autopolyploids display multivalent pairing and multisomic inheritance, whereas allopolyploids display bivalent pairing and disomic inheritance. According to genetic criteria, autopolyploids have more than two homologous genomes (i.e., AAAA) whereas allopolyploids have nonhomologous genomes (homoeologous) that arose from each species, (i.e., AABB). Given that the inheritance of many species has not been studied, estimating the number of auto and allopolyploids that exist in nature is challenging (Doyle and Broyles., 2017).

Autopolyploids can be produced through both genome doubling or unreduced gametes. Genome doubling can occur in many ways. When a normal diploid cell is (AA) subjected to heat shock, it can affect mitosis and produce a cell with double the chromosome number (AAAA) (Ramsey and Schemske., 1998). An autopolyploid can arise through intraspecific hybridization and subsequent genome doubling within a single individual (Soltis and Soltis, 2013, Corneillie et al., 2019). Experimentally, autopolyploids can also be produced with the use of colchicine.

Colchicine acts as a spindle inhibitor preventing regular disjunction of chromosomes; when a cell is treated with colchicine it will inhibit the cell division and produce one cell with double the sporophytic chromosome number (Briggs and Walters 1997).

Autopolyploids of different ploidies can be produced several different ways. Autotriploids are produced via crosses between gametes from a diploid and a tetraploid, or through the union of a reduced (1n) and unreduced gamete (2n). Triploid plants may backcross with their diploid progenitors and can occasionally produce tetraploids, which is known as triploid bridge. However, not all species can successfully form polyploids via a triploid bridge because of the triploid block. Autotetraploids can also be formed directly by merging two unreduced 2n gametes. For example, the union of two unreduced 2n gametes was the main mechanism of autotetraploid formation in *Anthoxanthum alpinum* (Bretagnolle 2001). Autohexaploids are produced in a tetraploid system by the union of reduced and unreduced gametes (Ramsey and Schemske 1998). Autopolyploidy is predominant in several angiosperm families, such as in Saxifragaceae and Cactaceae, whereas some families (e.g., Poaceae) have both allopolyploid and autopolyploid species (Soltis and Soltis 2013).

Allopolyploid species are derived from hybridization between two species. Allotetraploids can be formed directly from diploids after interspecific hybridization via genome doubling. Allotetraploids can also be formed through a triploid bridge. The F1 or F2 generation of interspecific crosses is frequently triploid, and self-fertilization or backcrossing of the triploids with diploids can sometime produce allotetraploids.

Several methods can be used to differentiate allopolyploids from autopolyploids. The traditional method is to examine pairing of chromosomes at meiosis. If the species is an allotetraploid, then only the homeologous chromosomes will pair at the first metaphase of

meiosis, producing a bivailent pattern similar to diploids (Darlington 1937; Stebbins 1945). If the individual is an autotetraploid, then the chromosomes can form multivailent structures because each chromosome will have more than one homolog with which to pair. Another method is to use molecular markers such as microsatellites to determine whether a polyploid exhibits disomic or multisomic inheritance patterns (Stift et al. 2008). Other studies have used other genetic approaches to determine whether polyploids exhibit genetic material from one or two potential diploid progenitors; for example, in the orchid *G. conopsea*, which contains both diploid and tetraploid individuals, they used the internal transcribed spacer (ITS) sequences to determine how the tetraploid formed. They concluded that the tetraploid individuals were likely autopolyploids because the only other species available with which to hybridize would be *G. densiflora,* yet the internal transcribed spacer (ITS) sequences differed between *G. conpsea* and *G. densiflora* but not between diploid and tetraploid *G. conopsea* plants (Jersáková et al., 2010).

1.4 Establishment of polyploids

To establish a population, newly formed polyploids must either colonize a new environment or must exist sympatry with its diploid progenitor. When the newly formed polyploid species has limited seed dispersal, then they must live sympatrically with their diploid progenitors (Hegarty and Hiscock, 2008, Levin 1975). It is challenging for a new polyploid species to establish in a mixed-cytotype population due to obstacles such as meiotic irregularities, reduced fertility, altered gene dosage, altered physiological properties, and minority cytotype disadvantage (Levin 2021). Genome doubling can give rise to changes in transpiration, water use efficiency, photosynthetic rate, phenology, and morphology in some polyploid species (Yves Van de peer et

al., 2021), which can sometimes have negative effects on fitness. A newly formed polyploid in a diploid population may also experience a frequency-dependent fitness disadvantage, as the new polyploid species may not have a compatible partner with whom to mate, termed "minority cytotype exclusion" (Levin 1975).

Newly formed polyploids overcome minority cytotype exclusion through different mechanisms. Recurrent formation of polyploids is an important mechanism providing polyploid individuals with compatible mates and at the same time generating genetic diversity in a newly formed polyploid species. Evidence suggests that most polyploid species have likely formed recurrently from different populations of their progenitors, followed by crossing between individuals of independent origin, thereby generating new genotypes (Soltis and Soltis 1999). For example, *Galax uroculata* has independently arisen 46 times, which is the greatest recorded number of independent polyploidization events for an auto or allopolyploid plant species (Levin 2021).

Assortative mating can also help polyploids overcome minority cytotype exclusion. Plants ensure assortative mating by spatial segregation of cytotypes via clonal expansion, niche differentiation, and limited seed dispersal. Asynchronous flowering and assortative pollen transfer are some other mechanisms to ensure the assortative mating (Kolar et al., 2017).
In other plants, polyploidization is often associated with a transition from cross-pollination to self-pollination, which enables them to overcome the problem of mating with their diploid progenitor. Divergence in traits attractive to pollinators may also help attract different pollinator communities, thereby promoting reproductive isolation in a mixed ploidy system (Van de peer et al., 2017).

In some cases, greater reproductive success can increase the ability of the polyploids to remain in a mixed-ploidy population and even displace their diploid progenitors. One main factor contributing to the successful establishment of polyploids is through the formation of large number of seeds, which is also a mechanism observed among invasive plants. Many studies have found that the production of many propagules in the initial phase of polyploid formation improves the likelihood of successful establishment (Levin 2021) because a larger founder population with more genetic diversity is likely to be more successful, as they may have greater adaptive potential.

1.5 Speciation via polyploidization

Because allopolyploids have chromosomes derived from two progenitor species, they may not be interfertile with their progenitors. In addition, they may be isolated through pre-mating barriers such self-fertilization, flowering time asynchrony, and pollinator-mediated isolation (Porturas and Segraves., 2019). Further, their populations may slowly increase because of mechanisms to overcome the minority cytotype disadvantage in a mixed ploidy population, such as recurrent formation of polyploids and self-pollination (Levin 1975, Barringer 2007). Because of physiological differences experienced by different cytotypes, polyploids may become ecologically differentiated on small scales. A shift in floral morphology may lead to pollinator-mediated assortative mating. All of these mechanisms may contribute to the speciation of polyploids, potentially helping to prevent their extinction (Husband and Sabara, 2003).

In contrast, many barriers may reduce the ability for new polyploid species to establish. For example, pollen-stigma incompatibility and pollen tube competition act as early post-mating barriers between polyploids and their progenitors, potentially preventing a polyploid from reproducing. When successful fertilization does occur, the unbalanced ratio of maternal and

paternal chromosomes in the endosperm may prevent the development of triploid (i.e., triploid block) (Portural and Segraves, 2019). In the absence of reproductive isolation, the offspring produced in crosses between the two cytotypes would be triploids with partial fertility, which often gives rise to gametes with an abnormal number of chromosomes. Even after establishment, if the polyploid is sympatric with their progenitor and has lower fitness, then it may have difficulty in establishing unless it disperses to a new environment (Soltis and Soltis 2009). Thus, because of these barriers, the frequency that polyploidization results in speciation is very low, as it was estimated to be between 15% and 24% of all instances of polyploidization (Levin 2021).

1.5.1 Reproductive isolation via morphological differences in cytotypes

Overall, previous studies have observed wide variation in the extent to which polyploids and their progenitors exhibit morphological differences, with important consequences for reproductive isolation. One study investigated the extent to which floral size and morphology, floral phenology, and floral scent differed between three co-occurring cytotypes from the genus *Gymnadenia* (Orchidaceae), including tetraploid and octoploid *G. conopsea* and tetraploid *G. densiflora* (Jerasakova et al., 2010). Overall, the cytotypes had similar flower color and morphology, and even though they exhibited small differences in floral scent and partial temporal segregation, they exhibited little evidence of pre-mating barriers, suggesting that other mechanisms may contribute to the cytotypes' coexistence. In another study, floral traits differed among diploid and tetraploid cytotypes of *Chamerion angustifolium* and tetraploids had greater pollinator visitation rates than the diploids, which reduced inter-cytotype mating, promoting reproductive isolation (Husband and Schemske, 2000, Husband and Sabara, 2016). In another study on *Trifolium*, the newly formed polyploids showed immediate changes in floral phenotype. Specifically, floral development was delayed in neopolyploids relative to the diploids, possibly

because the larger flower size takes more time to develop (Porturas and Segraves., 2019). Thus, polyploids exhibit wide variation in the extent to which they differ morphologically from their progenitors, which may affect their reproductive isolation.

1.5.2 Reproductive isolation by pollinators

Pollinators may also have an important role in mediating reproductive isolation between diploids and polyploids. For example, in *Gymnadenia conopsea,* tetraploids were visited by more pollinators and had a higher pollination efficiency. This resulted in higher reproductive success, which enabled them to persist in a mixed-ploidy population. Furthermore, in one population, they even eliminated their diploid counterpart (Gross and Schiestl, 2015). However, if polyploids arose from one or a small number of polyploidization events, pollinator-mediated reproductive success could be reduced by high inbreeding depression. In other species, divergence between polyploids and their progenitors was mediated by adaptive shifts in flowering phenology (Husband and Schemske, 2000, Gross and Schiestl, 2015).

In another study that focused on *Chamerion angustifolium*, the establishment of tetraploids was likely mediated by pollinator activity. For example, when the frequency of tetraploids was lower, pollinator visits were higher, whereas pollinator visitation rates were similar regardless of the frequency of diploids. Tetraploid reproductive success did not depend on their density, as they were successful regardless of whether they are isolated or rare (Gross and Schiestl, 2015). Therefore, pollinator activity has likely helped the newly formed tetraploids to establish in the mixed ploidy population, helping eliminate minority cytotype exclusion.

1.6 Adaptation and ecological consequences of polyploidy

It has long been observed that polyploid species occupy broader geographical extremes than their diploid progenitors and that they have greater capacity to tolerate severe conditions, such as cold

conditions in higher latitudes and elevations. Historical work on this phenomenon was pioneered by a Swedish cytologist Tackholm (1922) and later, other scientists also provided important reviews on this phenomenon (Rice at al., 2019). The fact that polyploids are generally more frequent in extreme environments has been attributed to increased copies of each chromosome, which increases genetic diversity (Corneillie et al., 2019). This increased genetic diversity will have a direct effect on gene expression, epigenetics, gene networks, cell size, and stress reactions (Fox et al., 2020).

Because of the increase in the DNA content, cell size is larger in polyploids, and this can affect growth and fitness, particularly for specific organs that have a determinant growth (Muntzing 1936, Corneielle *et al.,* 2019), which can have important phenotypic effects that may alter plant fitness and survival. For example, Corneillie et al. (2019) assessed the difference in phenology, growth, and cell wall composition among different ploidies (tetraploid, hexaploid and octoploid) of *Arabidopsis thaliana* plants. They observed a delay in phenology in plants with higher ploidies and a reduced cell division rate in the leaves, which is thought to compensate for increased cell size. Octoploids had a reduced cell division rate and therefore did not show an increase in biomass production, known as "high ploidy syndrome," which is thought to be due to an increased energy demand to support cell division (Corneillie et al., 2019).

The ecological and environmental consequences of polyploidy were recently highlighted in a recent study by Rice et al., 2019, who compiled geographic and lifeform databases with large-scale ploidy inference of plant species to infer the global distribution of polyploidy. Polyploid frequency, here referred to as the relative proportion of polyploid species out of all species with ploidy estimates, was greatest in the tundra biome in the far northern hemisphere. The Taiga biome had the second largest proportion of polyploid existence (47%) followed by the

temperate zones (38-40%) and montane grasslands (39%). The lowest percentages were found in the tropical and subtropical biomes. The greater frequency of polyploids in the polar latitudes was attributed to an increased frequency of unreduced gametes in the cold environments of the polar regions (Rice et al., 2019). They further analyzed the polyploid frequency at a finer scale by mapping it in the ecoregions of the world. However, the polyploid frequency of 39 ecoregions were markedly different from their biomes. A greater frequency of polyploids than expected given their biomes was found for Hawaii and the Andes, which are famous for being polyploid-rich. Two ecoregions, Montane fynbos and the renosterveld ecoregion, had a lower frequency compared to their biomes (Rice et al., 2019). Furthermore, unoccupied habitats such as recently deglaciated places and anthropogenically disturbed areas contained a greater frequency of polyploids. Polyploids were less abundant in areas with poor phosphorous and nitrogen availability, which could be due to the strong requirement for nucleic acids (Rice et al., 2019). Taxonomic composition was another factor affecting the frequency of polyploidy, with polyploids more commonly occurring in the regions with more commelinids and less rosids. Finally, they also observed that polyploidy is often associated with either self-compatibility or seed/vegetative apomixis.

Additionally, Rice et al. (2019) found that perennial herbs have greater levels of polyploidy (39%) than annual plants (28%), and woody plants (22%). The fact that a greater number of polyploids were perennial herbs may be attributable to their life span, which may provide them with a greater chance to find a suitable polyploid mate, helping to overcome the minority cytotype disadvantage. However, the frequency of polyploids among woody perennial plants was lower, which is counter to results observed for perennial herbs. Overall greater

frequencies of polyploids were observed in regions with lower woody species and greater perennial plant frequency (Rice et al., 2019).

1.7 Evolution of polyploidy within family Brassicaceae

Past polyploidization events have affected the diversification of almost all land plants, but it has played a prominent evolutionary role in some specific angiosperm lineages (Huang et al., 2020). For example, the Brassicaceae-specific palaeopolyploidization event was a duplication event that occurred approximately 50 million years ago and is shared among all Brassicaceae. Brassicaceae experienced accelerated diversification in the past 30 million years, which may be associated with this early polyploidization event (Huang et al., 2020). Brassicaceae has also experienced genome duplication events throughout its evolutionary history. Mesopolyploidization events (i.e. an event that occurred at intermediate periods in the evolutionary history of a group) have been documented for the tribes Anastaticeae, Biscutelleae, Brassiceae, Cochlearieae, Heliophileae, Iberideae, Schizopetaleae, Thelypodieae, Microlepideae, Physariae and Stevenieae (Huang et al., 2020), and the tribe Physariae is thought to be species-rich due to the occurrence of a whole-genome duplication event in its history (Hohmann et al., 2015). Frequent polyploidization (i.e., neopolyploidization), events have also occurred in the family, such that approximately 43.3% of the species in Brassicaceae are neopolyploids (Huang et al., 2020).

Brassicaceae is therefore a good model system to study ploidy-driven diversification of angiosperms, as it has a large number neopolyploid lineages that are potential sources for the future diversification.

Most WGD in Brassicaceae are associated with subsequent diploidization, which is achieved by eliminating large fractions of the duplicated genome over time. This process provides a platform for gene neofunctionalization, giving rise to new features that helped the

group adapt to the changing environment (Hohmann et al., 2015). For example, an herbivore, Pierid moths, evolved a glucosinolate detoxification mechanism around 34 million years ago (MYA), which led them to shift host plants to the Brassicales. The Brassicaceae experienced a whole-genome duplication event around the same time that is associated with the evolution of novel chemical defense compounds, the timing of which matches perfectly the onset of crown group radiation in Brassicaceae ~32 mya (Hohmann et al., 2015). Interestingly, one of the forefathers of polyploidy research, Stebbins (1950) proposed that polyploidy is unlikely to be a major force affecting diversification in major groups of plants, and instead proposed that polyploidization might drive the diversification of species and genera within families. However, studies on Brassicaceae have changed this perspective, as there is a whole genome duplication at the base of the family which likely accelerated diversification. However, researchers found no causal link between mesopolyploidization events and the diversification rates in Brassicaceae, as some of the WGD events are at the beginning of some species-rich tribes but others are found in comparatively small tribes (Huang et al., 2020).

1.7.1 Incidence of polyploidy within genus *Physaria*

One genus in the Brassicaceae family that exhibits wide variation within and among species in chromosome number and ploidy is *Physaria*. Fundamental chromosome numbers of the genus *Physaria* have been identified as x= 4, 5, 6, 7, 8 and 9 (Rollins 1939, Rollins 1966), and many polyploids have been identified with base chromosome numbers of x=5, x=6 and x=9 (Rollins 1966, 1971). Polyploids based on x= 6 are quite common and include *Physaria arenosa* (2n= 18); probably *Physaria argyrea* (n=18), *Physaria engelmanni* (n=18); *Physaria fendleri* (n=12), and *Physaria intermedia* (2n=18). Rollins (1971) also described polyploid populations based on n=5 such as *Physaria ludoviciana* and *Physaria arizonica*. *Physaria arizonica* has two

chromosome numbers (n=10 and 2n=10) indicating that polyploids occur in the species, and the diploid *Physaria arizonica* was found within 20 miles of the tetraploid. However, some other polyploid species in the genus show different chromosome numbers, such as *Physaria australis* (2n=14). Aneuploidy was also found to be common in the genus (Rollins and Rudenburg 1979); aneuploid series occur in both *Physaria argyraea* and *Physaria fendleri* (previous *Lesquerella*). (Rollins and Rudenburg 1979). The aneuploid population of *Physaria fendleri* has an n=7 chromosome count and was discovered only from New Mexico (Rollins 1977).

Several species of *Physaria* exhibit very complex patterns of chromosomal variation. The greatest range in chromosome number recorded in *Physaria* is in *Physaria argyraea.* The species also exhibits variation in morphology, leading Rollins to conclude that there might be more than one taxon whose boundaries are unclear, and therefore termed *Physaria argyraea* a species complex. The variation in chromosome numbers and morphology may also be due to hybridization (Rollins, 1966). Because of the difficulties in clarifying relationships in this species complex, *Physaria argyraea* was retained as a polymorphic species with two subspecies (Rollins 1977). *Physaria acutifolia* is another species showing a complex cytological situation, with the presence of several different chromosomal races, 2n = 8, 2n = 10, 2n = 16 and 2n =24. Rollins (1977) proposed that this species may hybridize with *Physaria chambersii* and suggested that *Physaria acutifolia* could be a possible agamic species (i.e., a species that includes a series of polyploid sexual, and facultatively or obligately agamospermous microspecies, which may result from hybridization among sexual diploid and sexual polyploid members of the complex). Although Rollins has proposed many scenarios involving hybridization and polyploidization, few of these hypotheses have been investigated using modern genetic approaches.

Many species in *Physaria* were proposed to be mixed-cytotype autopolyploids, including *P. alpina, P. arenosa, P. arizonica, P. argyraea, P. calcicola, P. fendleri, P. intremedia, P. ludoviciana, P. ovalifolia, P. purpurea, P. rectipes, and P tenella* (Rollins and Rudenburg 1979). However, in these species, previous researchers found no morphological, physiological and ecological differences between individuals of different ploidy. Even though it is likely that chromosomal differences act as a barrier to gene flow between individuals of different ploidy, Rollins and Shaw (1973) found little justification to recognize them as different species. However, if polyploidization produces taxonomically important traits that can be used to distinguish the ploidies, then Rollins suggested that the polyploids should be identified as taxonomically different species. Although Rollins suggested that both situations occur in the genus *Physaria*, no study has investigated the origins of these species with genetic approaches.

1.8 Future directions

Polyploidy has been studied extensively in only a few dominant plant groups; it is therefore important to incorporate more complete representation of species from a more diverse assortment of plant groups. Additional work is also needed to incorporate additional estimates of ploidy into floristic work, particularly in groups that are known to include polyploids, as it is likely that estimates of number of known polyploid taxa are incomplete. This is particularly true in the global south. In addition, additional studies need to be conducted to assess the factors influencing the establishment of polyploids in a diploid population, for example; by conducting studies into how polyploidy affects floral traits, and how that in turn affects reproductive isolation in mixed-ploidy populations. Finally, additional work is needed to understand the evolutionary origins, reproductive isolation, and establishment of polyploids in specific plant groups; for example,

research is needed to understand ecology, evolution, and reproductive dynamics of polyploids in groups exhibiting large variation in ploidy such as the genus *Physaria*.

GLOSSARY

Polyploidy – Having more than two chromosome sets per nucleus in an organism.

Autopolyploid – A polyploid species arise through interspecific hybridization and subsequent genome doubling within a single individual or self-fertilization involving unreduced gametes

Allopolyploid – Allopolyploid species derive from the crosses between two species.

Neo polyploid – A newly formed polyploid within a mixed ploidy population.

Cytotype – A chromosomal variant within a species. Eg: 2x(diploid), 3x (triploid), 4x (tetraploid) and higher ploidies of a same species.

Mixed ploidy – Having more than one cytotype within a population eg: diploid and tetraploid plants in the same population

Minority cytotype exclusion – Elimination of the minor/rare cytotype from a diploid population, because the minor cytotype face a frequency dependent fitness disadvantage and by mating with the dominant diploid cytotype fewer fit offspring will be produced.

Unreduced gametes- a pollen or egg cell nuclei having somatic (2n) chromosome number.

Triploid block – The inability to form viable seeds of triploid cytotype due to the prevention of endosperm development due to the unbalanced maternal and paternal chromosome ratio.

Triploid bridge – Formation of tetraploids or higher ploidies via the fertile triploid cytotype backcrossed to a diploid individual.

REFERENCES

Barringer, B. C. (2007). Polyploidy and self-fertilization in flowering plants. *American Journal of Botany*, *94*(9), 1527–1533. https://doi.org/10.3732/ajb.94.9.1527

Bretagnolle, F. (2001). Pollen production and spontaneous polyploidization in diploid populations of Anthoxanthum alpinum. *Biological Journal of the Linnean Society*, *72*(2), 241–247. https://doi.org/10.1111/j.1095-8312.2001.tb01314.x

Bretagnolle, F., & Thompson, J. D. (1995). Gametes with the somatic chromosome number: Mechanisms of their formation and role in the evolution of autopolyploid plants. *New Phytologist*, *129*(1), 1–22. https://doi.org/10.1111/j.1469-8137.1995.tb03005.x

Briggs, D. and Walters, S.M. (1997) 'Abrupt speciation', in *Plant Variation and Evolution*. Cambridge: Cambridge University Press.

Castro, S., Loureiro, J., Procházka, T., & Münzbergová, Z. (2012). Cytotype distribution at a diploid–hexaploid contact zone in *Aster amellus* (Asteraceae). *Annals of Botany*, *110*(5), 1047–1055. https://doi.org/10.1093/aob/mcs177

Corneillie, S., De Storme, N., Van Acker, R., Fangel, J. U., De Bruyne, M., De Rycke, R., Geelen, D., Willats, W. G. T., Vanholme, B., & Boerjan, W. (2019). Polyploidy Affects Plant Growth and Alters Cell Wall Composition. *Plant Physiology*, *179*(1), 74–87. https://doi.org/10.1104/pp.18.00967

Darlington, C.D. 1937. Recent advances in cytology. Philadelphia:Blakiston

Doyle, J. J., & Sherman-Broyles, S. (2017). Double trouble: Taxonomy and definitions of polyploidy. *New Phytologist*, *213*(2), 487–493. https://doi.org/10.1111/nph.14276

Fox, D. T., Soltis, D. E., Soltis, P. S., Ashman, T.-L., & Van de Peer, Y. (2020). Polyploidy: A Biological Force From Cells to Ecosystems. *Trends in Cell Biology*, *30*(9), 688–694. https://doi.org/10.1016/j.tcb.2020.06.006

Gross, K., & Schiestl, F. P. (2015). Are tetraploids more successful? Floral signals, reproductive success and floral isolation in mixed-ploidy populations of a terrestrial orchid. *Annals of Botany*, *115*(2), 263–273. https://doi.org/10.1093/aob/mcu244

Hegarty, M. J., & Hiscock, S. J. (2008). Genomic Clues to the Evolutionary Success of Polyploid Plants. *Current Biology*, *18*(10), R435–R444. https://doi.org/10.1016/j.cub.2008.03.043

Heslop-Harrison, J. S. (Pat), Schwarzacher, T., & Liu, Q. (2023). Polyploidy: Its consequences and enabling role in plant diversification and evolution. *Annals of Botany*, *131*(1), 1–10. https://doi.org/10.1093/aob/mcac132

Hohmann, N., Wolf, E. M., Lysak, M. A., & Koch, M. A. (2015). A Time-Calibrated Road Map of Brassicaceae Species Radiation and Evolutionary History. The Plant Cell, tpc. 15.00482. https://doi.org/10.1105/tpc.15.00482

Huang, X.-C., German, D. A., & Koch, M. A. (2020). Temporal patterns of diversification in Brassicaceae demonstrate decoupling of rate shifts and mesopolyploidization events. *Annals of Botany*, *125*(1), 29–47. https://doi.org/10.1093/aob/mcz123

Husband And, B. C., & Schemske, D. W. (2000). Ecological mechanisms of reproductive isolation between diploid and tetraploid *Chamerion angustifolium*. Journal of Ecology, 88(4), 689–701. https://doi.org/10.1046/j.1365-2745.2000.00481.x

Husband, B. C., & Sabara, H. A. (2004). Reproductive isolation between autotetraploids and their diploid progenitors in fireweed, *Chamerion angustifolium* (Onagraceae). *New Phytologist*, *161*(3), 703–713. https://doi.org/10.1046/j.1469-8137.2004.00998.x

Husband, B. C., Baldwin, S. J., & Sabara, H. A. (2016). Direct vs. indirect effects of whole-genome duplication on prezygotic isolation in *Chamerion angustifolium*: Implications for rapid speciation. *American Journal of Botany*, *103*(7), 1259–1271. https://doi.org/10.3732/ajb.1600097

Husband, B. C., Baldwin, S. J., & Suda, J. (2013). The Incidence of Polyploidy in Natural Plant Populations: Major Patterns and Evolutionary Processes. In J. Greilhuber, J. Dolezel, & J. F. Wendel (Eds.), *Plant Genome Diversity Volume 2* (pp. 255–276). Springer Vienna. https://doi.org/10.1007/978-3-7091-1160-4_16

Jersáková, J., Castro, S., Sonk, N., Milchreit, K., Schödelbauerová, I., Tolasch, T., & Dötterl, S. (2010). Absence of pollinator-mediated premating barriers in mixed-ploidy populations of *Gymnadenia conopsea* s.l. (Orchidaceae). *Evolutionary Ecology*, *24*(5), 1199–1218. https://doi.org/10.1007/s10682-010-9356-7

Kolář, F., Čertner, M., Suda, J., Schönswetter, P., & Husband, B. C. (2017). Mixed-Ploidy Species: Progress and Opportunities in Polyploid Research. *Trends in Plant Science*, *22*(12), 1041–1055. https://doi.org/10.1016/j.tplants.2017.09.011

Levin, D. A. (1975). Minority cytotype exclusion in local plant populations. *Taxon*, *24*(1), 35–43. https://doi.org/10.2307/1218997

Levin, D. A. (2021). Propagule pressure and the establishment of emergent polyploid populations. *Annals of Botany*, *127*(1), 1–5. https://doi.org/10.1093/aob/mcaa187

Müntzing, A. (1937). The Evolutionary Significance of Autopolyploidy. Hereditas, 21(2–3), 363–378. https://doi.org/10.1111/j.1601-5223.1936.tb03204.x

Oleszczuk, S., Grzechnik, N., Mason, A. S., & Zimny, J. (2019). Heritability of meiotic restitution and fertility restoration in haploid triticale. *Plant Cell Reports*, *38*(12), 1515–1525. https://doi.org/10.1007/s00299-019-02462-6

Porturas, L. D., & Segraves, K. A. (2020). Whole genome duplication does not promote common modes of reproductive isolation in *Trifolium pratense*. *American Journal of Botany*, *107*(5), 833–841. https://doi.org/10.1002/ajb2.1466

Ramanna, M. S., & Jacobsen, E. (n.d.). *Relevance of sexual polyploidization for crop improvement – A review*.

Ramsey, J., & Schemske, D. W. (1998). Pathways, mechanisms, and rates of polyploid formation in flowering plants. *Annual Review of Ecology and Systematics*, *29*(1), 467–501. https://doi.org/10.1146/annurev.ecolsys.29.1.467

Rice, A., Šmarda, P., Novosolov, M., Drori, M., Glick, L., Sabath, N., Meiri, S., Belmaker, J., & Mayrose, I. (2019). The global biogeography of polyploid plants. *Nature Ecology & Evolution*, *3*(2), 265–273. https://doi.org/10.1038/s41559-018-0787-9

Rollins, R. C. (1939). The cruciferous genus *Physaria*. Rhodora 41: 391-414.

Rollins, R. C. (1966). Chromosome numbers of Cruciferae. *Contributions from the Gray Herbarium of Harvard University*, *no. 192-197 1963-66*(197), 43–65. https://doi.org/10.5962/p.336417

Rollins, R. C., & Rüdenberg, L. (1971). Chromosome numbers of Cruciferae. II. *Contributions from the Gray Herbarium of Harvard University*, *no. 198-202 1969-71*(201), 117–133. https://doi.org/10.5962/p.336429

Rollins, R. C., & Rüdenberg, L. (1977). Chromosome numbers of Cruciferae III. *Contributions from the Gray Herbarium of Harvard University*, *no. 206-208 1976-78*(207), 101–116. https://doi.org/10.5962/p.336444

Rollins, R. C., & Rüdenberg, L. (1979). Chromosome numbers of Cruciferae. IV. Contributions from the Gray Herbarium of Harvard University.

Rollins, R., & Shaw, E. (1973). The Genus *Lesquerella* (Cruciferae) in North America. (pp. 4-47). Cambridge, MA: Harvard University Press.

Schinkel, C. C. F., Kirchheimer, B., Dullinger, S., Geelen, D., De Storme, N., & Hörandl, E. (2017). Pathways to polyploidy: Indications of a female triploid bridge in the alpine species Ranunculus kuepferi (Ranunculaceae). *Plant Systematics and Evolution*, *303*(8), 1093–1108. https://doi.org/10.1007/s00606-017-1435-6

Soltis, D. E., & Soltis, P. S. (1999). Polyploidy: Recurrent formation and genome evolution. *Trends in Ecology & Evolution*, *14*(9), 348–352. https://doi.org/10.1016/S0169-5347(99)01638-9

Soltis, D. E., Visger, C. J., & Soltis, P. S. (2014). The polyploidy revolution then...and now: Stebbins revisited. *American Journal of Botany*, *101*(7), 1057–1078. https://doi.org/10.3732/ajb.1400178

Soltis, P. S., & Soltis, D. E. (2009). The Role of Hybridization in Plant Speciation. *Annual Review of Plant Biology*, *60*(1), 561–588. https://doi.org/10.1146/annurev.arplant.043008.092039

Sora, D., Kron, P., & Husband, B. C. (2016). Genetic and environmental determinants of unreduced gamete production in Brassica napus, Sinapis arvensis and their hybrids. *Heredity*, *117*(6), 440–448. https://doi.org/10.1038/hdy.2016.69

Stebbins, G. L. (1947). Types of polyploids: Their classification and significance. *Advances in Genetics*, 403–429. https://doi.org/10.1016/s0065-2660(08)60490-3

Stift, M., Berenos, C., Kuperus, P., & Van Tienderen, P. H. (2008). Segregation Models for Disomic, Tetrasomic and Intermediate Inheritance in Tetraploids: A General Procedure Applied to Rorippa (Yellow Cress) Microsatellite Data. *Genetics*, *179*(4), 2113–2123. https://doi.org/10.1534/genetics.107.085027

Toda, E., & Okamoto, T. (2020). Polyspermy in angiosperms: Its contribution to polyploid formation and speciation. *Molecular Reproduction and Development*, *87*(3), 374–379. https://doi.org/10.1002/mrd.23295

Van de Peer, Y., Ashman, T.-L., Soltis, P. S., & Soltis, D. E. (2021). Polyploidy: An evolutionary and ecological force in stressful times. *The Plant Cell*, *33*(1), 11–26. https://doi.org/10.1093/plcell/koaa015

Van de Peer, Y., Mizrachi, E., & Marchal, K. (2017). The evolutionary significance of polyploidy. *Nature Reviews Genetics*, *18*(7), 411–424. https://doi.org/10.1038/nrg.2017.26

Wood, T. E., Takebayashi, N., Barker, M. S., Mayrose, I., Greenspoon, P. B., & Rieseberg, L. H. (2009). The frequency of polyploid speciation in vascular plants. *Proceedings of the National Academy of Sciences*, *106*(33), 13875–13879. https://doi.org/10.1073/pnas.0811575106

CHAPTER 2 PHYLOGENY AND TAXONOMY OF GENUS *PHYSARIA* IN NORTH AMERICA

2.1 Introduction

Reconstructing the phylogeny of a large group of plants can be challenging due to issues such as hybridization, polyploidy, and incomplete lineage sorting. However, reconstructing the phylogeny of a group is important because it provides an evolutionary framework to elucidate the evolution of morphological characters, delimit species, identify hybrids, clarify taxonomy, and provide information that is relevant for conservation (Beardsley et al., 2004). For example, phylogenetic information can help clarify whether species form monophyletic groups, which is important because many species concepts rely on the monophyly of the species (de Queiroz and Donoghue 1988). Phylogenies can help clarify the limits and status of a putative endangered species to understand whether it is a subpopulation of a more widespread species or instead whether it represents a unique, endangered species that merits conservation. Given the limited resources available for conservation, evolutionary information is important because it can be used to prioritize the species that need to be conserved (Veron et al., 2019).

One group in which understanding the evolutionary history has been challenging due to its relatively recent radiation, possible incomplete lineage sorting, hybridization, and variation in chromosome number and ploidy is the genus *Physaria*. *Physaria* is a genus of ~108 predominantly new world plant species, the most species-rich genus in tribe Physariae, which also includes six additional genera: *Dimorphocarpa* (4 spp.), *Nerisyrenia* (7-11 spp.), *Lyrocarpa* (3 spp.), *Synthlipsis* (2 spp.), *Dithyrea* (2 spp.), and *Paysonia* (8 spp.). *Physaria* species generally form dense trichomes, which are thought to be an adaptation to the xeric, open environments preferred by the species (Mazie and Baum, 2016). Species in the genus *Physaria*

species are distinctive in having stellate trichomes whereas the rest of the family has branching trichomes (Mazie and Baum, 2016). *Physaria* species are either annuals, biennials or perennials, and have erect stems with a caudex often present. Flowers are usually yellow, purple, or rarely white (Flora of North America, vol:7).

Despite all the features that unite *Physaria* species, the genus exhibits wide variation in many characteristics, such as chromosome number. Several studies have documented variation in chromosome numbers in *Physaria* species and have speculated about the base chromosome in the genus (Rollins 1939, Mulligen 1966, Rollins 1973). Species vary in chromosome number; for example, chromosome numbers in *Physaria* can vary from n=4, or 5 in several related species in a "low chromosome numbers" clade (Fuentes Soriano and Kellogg 2021) to 23 in polyploid individuals of *Physaria* (Salywon et al. 2022). Species also exhibit intraspecific variation in chromosome number; for example, *Physaria fendleri* exhibits variation in chromosome number from 2*n* = 12, 14, 24, which includes both diploids and putative tetraploids. Chromosome number is identified as one mechanism responsible for diversification at both the intraspecific and species level (Wood et al., 2009), and complex genomic processes such as polyploidy have been proposed as a factor that may have shaped the diversification and adaptation of *Physaria* species in the higher montane zones (Fuentes-Soriano and Kellogg., 2021). Chromosome numbers have been an important character in making taxonomic decisions for some *Physaria* species but not for others (Rollins and Rudenberg, 1979). The evolution of chromosome number and polyploidy in *Physaria* is still poorly known because a well-resolved phylogeny that includes a large proportion of the North American species was lacking.

Most *Physaria* species are endemic to North America (Flora of North America, vol:7). *Physaria* species occur across western North America, from the southwestern United States to the Canadian border and from the Great Plains to the Sierra Nevada and cascade mountain ranges (Rollins 1939, Waite 1973, Rollins, and Shaw 1973). The center of diversity of *Physaria* occurs in the arid regions of the great basin, between the ranges of Rocky Mountains, and in the southwestern United states. However, several species occur in South America, and one species, *Physaria arctica*, extends into Greenland, northern Canada, Alaska and arctic Russia (Rollins, 1939, Rollins, 1981). The geographic range size of species in the genus *Physaria* also varies widely, with the ranges of some species extending across most of western North America, whereas many species have narrowly endemic ranges that may extend only across a very small geographic area (Rollins and Shaw 1973, Rollins 1993). For example, many species in Mexico are narrowly endemic (*P. argentea, P. inflata, P. mexicana, P. mirandiana, P. rosei, P. wyndii*; Fuentes Soriano and Kellogg., 2021). Some *Physaria* species with endemic distributions are listed as endangered or threatened as well, therefore it is important to prioritize conservation actions for them. However, many endemic species are taxonomically questionable, some species may hybridize, and some may have undergone both allo- or auto-polyploidization, leading to questions about the distinctiveness of many species. Understanding the limits and status of species is important for implementing effective conservation strategies (Veron et al., 2019).

The wide variation in characters such as floral traits, geographic range size and chromosome number has overall led to conflicting interpretations of species relationships in the group. Some relationships among species in the group were proposed in monographs by Payson (1921) or Rollins and Shaw (1973), who suggested groups of species based on morphological characters such as pod shape, branching type of trichomes, and number of ovules (Payson 1921;

Rollins and Shaw, 1973). Payson grouped species into the Sections *Alyssum, Enantiocarpa* and *Eulesqurella* with one, three and forty-eight species, but only species in section *Eulesquerella* are currently placed in the genus *Physaria*. Payson also proposed many informal species groups (summarized in Table 2.1.). Subsequently, Rollins and Shaw (1973) grouped species into 10 informal groups that differed from those proposed by Payson (Table 2.1); whether either of these grouping strategies corresponds with clades is unclear.

The most recent and most comprehensive phylogenetic treatment of the genus to date, which aimed to resolve the relationships among the genera in tribe Physariae, included 29 of the 108 *Physaria* species and reconstructed the phylogeny using two nuclear markers, ITS and LUMINIDEPENDENS. The study resolved two strongly supported clades: the DDNLS clade, comprised of *Synthilpsis*, *Nerisyrenia*, *Dimorphocarpa*, *Lyrocarpha* and *Dithyrea*, and the PP clade, comprised of *Paysonia* and *Physaria*, supporting the previous studies on tribe Physarieae which identified two main clades (Fuentes-Soriano and Kellogg, 2021, Mazie and Baum 2016). The monophyly of both *Physaria* and *Paysonia* was supported, but the branches supporting these clades received variable support depending on the method used for phylogeny reconstruction (Fuentes-Soriano and Al-Shehbaz, 2013, Fuentes Soriano and Kellogg, 2021). Thus, a phylogeny that includes both sampling of a greater proportion of the 108 species of *Physaria* and a larger number of polymorphic loci is needed to understand the monophyly of species and implications for taxonomy, the evolutionary relationships among species, and the biogeography of the genus.

Here, our goal was to reconstruct the phylogeny of *Physaria*, including a large proportion of species (~86 /108 spp.) in the genus, to clarify the evolutionary history of the group. Based on

the resulting phylogeny, our goals are to 1) test the relationships among clades and species relationships proposed by Payson or Rollins and Shaw (summarized in Table 2.1), 2) to investigate the monophyly of species to aid in species delimitation, and 3) to investigate how the inclusion of polyploid taxa affects the topology of the phylogeny, which may help shed light on the origins of polyploid taxa.

2. 2 Materials and Methods

2.2.1 Study species and taxonomic history of *Physaria*.

Taxonomic history of Physaria. Initially, *Physaria* was first described by Torrey and Gray (1838) as a section of old-world genus *Vesicaria* (Rollins, 1939a, Al Shehbaz and O'Kane, 2002). In 1848, Asa Gray recognized *Physaria* as a separate genus comprised of New World, double-fruited species (Al Shehbaz and O'Kane, 2002). Watson (1888) recognized *Lesquerella* as a separate genus comprising species with non-inflated, non-double fruits (Al-Shehbaz and O'Kane, 2002).

There have been several taxonomic revisions on the genus *Physaria* (Payson 1921, Rollins and Shaw 1973, Rollins 1993, Al-Shehbaz and O'Kane 2002, Grady and O'Kane 2007). In the first monograph of genus *Lesquerella*, Payson first drew attention to similarities between species of genus *Physaria* and *Lesquerella* (Payson 1921). Payson grouped 52 species of *Lesquerella* into three sections (*Alyssum, Enantiocarpa* and *Eulesquerella*) based on two morphological characters, trichome type and inflorescence type, but only species in section *Eulesquerella* are currently placed in the genus *Physaria*. Payson also proposed the center of the origin of *Lesquerella* (now *Physaria*) to be Central Texas based on the number of species (*P. auriculata, P. grandiflora, P. densiflora, P. engelmannii, P. ovalifolia, P.argyrea, P. fendleri, P. recurvata, P.gracilis,*and *P.gordonii)* and the primitive characters of the species found in the region. The morphological characters that were proposed to be primitive in the genus are large

siliques, a large number of ovules dispersed entirely around the replum, and a predominance of simple trichomes compared to derived characters, which were thought to be small silliques, a small number of ovules concentrated around the upper replum area and differentiated trichomes (Payson 1921). He observed that the characters of the species diverge with distance from the proposed center of the origin, with the geographical extremes occupied by the most specialized species. For example, he proposed that the species occurring in South America and arctic America are very specialized and that they may have originated via long-distance dispersal from the center of distribution (Payson 1921).

In later monographs, Reed Rollins monographed 69 *Lesquerella* species (Rollins, 1973), and thereafter included 22 species in *Physaria* and 83 species in *Lesquerella* (Rollins, 1993). Subsequently, Rollins and Shaw grouped species into 10 informal groups that differed from those proposed by Payson (Table 2.1). They used combinations of characters including trichome patterns, pollen morphology, fruit morphology and chromosome counts and mentioned that the same character combinations would not work for all the species delimitations. Rollins also noted continuous morphological gradation from *Physaria* to *Lesquerella* (Rollins and Shaw, 1973). Rollins and Shaw proposed that because *Physaria* is most like *Alyssum*, which occurs in Yukon region in northwestern Canada and adjacent Alaska, that *Physaria* also may have possible migrated using the same route from Asia to North America (Rollins and Shaw,1973).

Subsequently, additional comparative studies in *Physaria* and *Lesquerella* showed similarities in habit, ecology, leaf morphology, trichome type, inflorescence, flower color, fruiting pedicels, seed coat sculpture, embryo type, and biogeography except for a set of auriculate-leaved species of *Lesquerella* (*Paysonia*) endemic to southeastern and central United States (O'Kane et al., 1999). Molecular studies in the genus revealed that *Physaria* was nested

within *Lesquerella* and evolved more than once from *Lesquerella* (O'Kane et al., 1999). Thus, based on molecular, morphological, and geographic data, most *Lesquerella* were transferred into *Physaria* except the 8 auriculate-leaved species, which were transferred to the genus *Paysonia* (Al Shehbaz and O' Kane, 2002). Altogether in 2002, 99 species of *Physaria* species were recorded (Al Shehbaz and O'Kane, 2002), but several additional species have since been described (e.g., Grady and O'Kane, 2007). The genus *Physaria*, as currently circumscribed, now contains ~108 species (Flora of North America Vol: 7)

Sample collection and DNA extraction. Of ~108 species of *Physaria* identified to date, a total of 222 samples representing 82 species were collected by Professor Steve O' Kane of University of Northern Iowa. Collection information is given in appendix 2.1. A sample of *Paysonia lasiocarpa* subsp. *berlandieri* was also collected to serve as the outgroup.

All lab work was conducted in the conservation genetics lab at Missouri Botanical Garden. DNA was extracted using a modified CTAB protocol with an additional two sorbitol washes (Doyle and Doyle, 1987, Storchova et al., 2000). The DNA concentration of each sample was quantified using a Qubit^{TM} fluorometer (Thermofisher). Samples with >500 ngul^{-1} was used for reduced-representation genomic library preparation for Illumina DNA sequencing.

2b-RAD-seq (Wang *et al*., 2012) was used to generate high-quality SNP data across the genome. We chose this approach because previous analyses that employed Sanger sequencing of plastid DNA regions was unable to resolve relationships (Steve O'Kane, pers. Comm.), suggesting that the genus represents a rapid radiation and that an approach that involves a larger number of polymorphic loci will be necessary to resolve relationships. The RAD-seq approach chosen for the study is very cost-effective and has been used to successfully resolve a rapid radiation within

a similarly sized genus (Acha et al., 2021). Quantified DNA was digested with the type II b restriction enzyme BcgI (New England Biolabs) to generate fragments of DNA from throughout the genome. The digestion was conducted in a 96 well plate and then unique double stranded adapters were ligated to samples in each of the 12 columns. Ligated DNA samples were amplified using High Fidelity Phusion PCR mix (New England Biolabs) for 14 PCR cycles. Amplified DNA was checked for the success of digestion and ligation using agarose gel electrophoresis before proceeding to the large-scale amplification. For the large-scale amplification, each uniquely barcoded sample in a row was pooled and amplified using one of eight uniquely barcoded PCR primers. This produced 96 uniquely barcoded samples per plate. The final PCR was amplified for 15 cycles with the same conditions as the test amplification and then checked using agarose gel electrophoresis. The resulting bands were excised and purified using MinElute Gel Extraction Kit (QIAGEN). The amount of DNA in each excised band was quantified using a Qubit fluorometer and pooled into a single pool with a final concentration of 10nM. The final pooled DNA samples were sent to Northwestern University and sequenced in 2.5 lanes on an Illumina HiSeq 4000 using single-end, 50 bp reads.

2.3 Data analysis

2.3.1 Sequencing quality control, assembly of loci and SNP calling.

The resulting sequences were assessed for quality using FastQC (Barraham Bioinformatics) (Linan et al., 2019, Acha et al., 2021). The sequences were demultiplexed using a 2bRAD de novo script written by M. Matz, which also removed barcodes and Illumina adapters (Linan et al., 2019, Acha et al., 2021). FastX toolkit was used to remove low-quality reads.

Sequences were aligned de novo and single nucleotide polymorphisms (SNPs) were identified using iPyrad v0.9.84 (Eaton and Overcast., 2020) with the parameters described in supplementary data table S1. For the assembly, the parameters were kept as the defaults except we set parameter 11 (minimum depth for statistical base calling) to 10, parameter 12 (minimum depth for the majority rule base calling) to 4, parameter 13 (maximum cluster depth within samples) to 8000, parameter 14 (clustering threshold for de novo assembly) to 0.9, and parameter 21 (minimum number of samples per locus) to 3% of the number of samples.

2.3.2 Phylogenetic analysis

Maximum likelihood (ML)

All phylogenetic trees were produced using the software IQ tree v2.2.0.3 (Nguyen et al., 2015) using the same parameters for all the datasets, which were: -m MFP which uses model finder plus to determine the best fit model for the dataset (Kalyanamoorthy *et al*., 2017). Model finder computes the Bayesian information criterion (BIC) scores for each model in the database to determine the best fit model and then runs the rest of the analysis using the best model for the Maximum Likelihood phylogeny reconstruction (Nguyen *et al*., 2015). We then ran IQ tree for 1000 bootstrap replicates. This bootstrap analysis uses the built-in ultrafast bootstrap approximation (UF Boot), which has been shown to result in similar branch support values as non-parametric traditional bootstrap but more quickly (Ming *et al*., 2013, Hoang *et al*., 2018).

Because the large number of samples and loci in the study made phylogeny reconstruction very computationally expensive, we first reconstructed a phylogeny including one sample per taxon (including species, subspecies, and variety) to gain an understanding of the main clades in the group, with a goal of conducting a sub-analysis separately for the sub-clades that included

multiple samples per taxon. Therefore, one sample per species was included in the first data set, totaling 100 samples with 82 species represented; samples were selected by choosing the one for each taxon with the highest sequencing depth, regardless of the percentage of missing data. The best model for this data set was chosen as TVM+F+I+I+R2 according to the BIC scores (TVM-Transversion model, AG=CT and unequal base freq).

ML analysis that included one sample per taxon recovered two geographically structured clades, hereafter referred to as the eastern and western clades (see Results). We then conducted a sub-analysis of each clade that included all samples of each taxon grouped in the clade. We included *Paysonia lasiocarpa* subsp. *berlandieri* as the outgroup for each clade. For this analysis, we conducted a new assembly for each clade and removed any sample that showed >95% missing data. The TVM+F+R2 model was selected as the optimal model of evolution for eastern clade whereas the TVM+F+I+G4 model of evolution was selected for the western clade. In total, the analysis of the eastern clade contained 24 samples representing 10 species and the western clade contained 84 samples representing 46 species.

2.3.3 Analysis of polyploid *Physaria* species

To understand how the inclusion of putative polyploid taxa affect the topology of the phylogenies, we first identified potential polyploid species of genus *Physaria* using the available literature. A total of 21 species were recorded as potential polyploids (Table 2.2). These potential polyploids were removed from the eastern and western clade data sets and then we re-analyzed the phylogeny of the two data sets.

2.4 Results

2.4.1 Locus assembly, SNP calling, and results of phylogeny reconstructions.

After initial inspection of sequencing depths, we removed samples that showed low sequence depths, resulting in a data set that included 212 accessions, representing one outgroup, 80 *Physaria* species and 20 subspecies of ingroup taxa.

Species tree including one sample per species. Our first data set included one sample per species and subspecies with the intent of placing every species in the phylogeny into major clades. The data set included 100 samples representing 80 species and 20 subspecific taxa. The total number of loci in the assembly was 53,586, of which 44,567 were variant sites. Samples ranged from 328 loci in the assembly to 15,911 with an average of 3,999 loci per sample, such that samples ranged from 70.3% to 99.4% missing data in the assembly, with an average of 92.5% missing data across all samples. Using the sample of *Paysonia* as the root, the phylogeny divided accessions into two main clades. Both clades had samples from Mexico and Texas at the base of the tree, but one clade contained predominately eastern North American species and the other clade contained predominately western North American species (Fig. 2.1). In general, all subspecies of a species were placed in the same clade except *P. argyraea*, in which *subsp. argyraea* was placed in the eastern clade and *subsp. diffusa* was placed in the eastern clade. Separate phylogenies were reconstructed for each of the resulting two clades, which are described below.

Eastern clade -This data set contained 25 samples representing 10 species and 4 subspecies of *Physaria* and one outgroup, including putative polyploids *P. argyraea* subsp. *argyraea* and *P.*

engelmannii. The total number of loci in the assembly was 47,280, of which 29,167 were variant sites. The number of loci in the assembly per sample varied from 2102 to 30,999, with an average of 12,735 loci across samples. Thus, the proportion of missing data per sample ranged from 34.4% to 96.8% with an average of 73.1% missing data per sample.

In the resulting phylogenetic tree, a strongly supported clade that contained two accessions of *P. argyraea* subsp. *argyraea* from Mexico and Texas (TX) and one accession of P. *densiflora* subsp. *densiflora* from Mexico was placed as sister to all other accessions followed by a grade made up of: 1) one accession of *P. engelmannii* from TX, 2) one accession of *P. densiflora* from TX, 3) clade A1 (100% bootstrap support [BS]) containing both of *P. recurvata* accessions from TX, 4) clade A2 (100% BS support) containing *P. gracilis* subsp. *nuttallii* from Oklahoma (OK) and *P. gracilis* subsp. *gracilis* from TX, 5) clade A3 (100% BS support) containing three *P. gracilis*, and 6) one accession of *P. sessilis* from TX. The remaining accessions formed a strongly supported clade (Clade B) (90% BS) that was further divided two subclades B1 and B2. Subclade B1 contained one accession of *P. argyraea* subsp. *argyraea* placed as sister to a strongly supported (100% BS) clade of three accessions of *P. angustifolia* from OK. Subclade B2 was divided into two groups, B2 (i) contained one accession of *P. filiformis* from MO placed as sister to a strongly supported (100% BS) monophyletic group of two accessions of *P. globosa* from Tennessee (TN), whereas B2 (ii) comprised a strongly supported (100% BS) monophyletic group of all four accessions of *P. ouachitensis* from Arkansas (Fig. 2.2). Of the 7 species that were represented by multiple accessions, *P. angustifolia, P. ouachitensis, P. recurvata, and P. globosa* were supported as monophyletic, whereas *P. gracilis*, *P. argyraea*, and *P. densiflora* were non-monophyletic.

Eastern clade without potential polyploid species - In this dataset, we removed four accessions of two putative polyploid species in the clade, such that 21 samples remained. The resulting assembly contained 75,590 loci, of which 44,343 were variant sites. The number of loci in the assembly varied from 1,970 to 45,514 loci, with an average of 20,684 loci. The resulting proportion of missing data range from 39.8% 97.4% with 72.6% average of missing data.

When we removed the putative polyploids *P. argyraea* subsp. *argyraea* and *P. engelmannii* (Table 2.2), the monophyly of species remained unchanged but the topology of the tree changed (Figs 2.3 and 2.4). *P. densiflora* subsp. *densiflora* from Mexico and *P. densiflora* from TX became successive sisters to the remainder of the clade, followed by the clade A1 containing three accessions of *P. gracilis* from TX. Next, *P. sessilis* from TX was placed as the sister group to clade A2 with two accessions of *P. recurvata* from TX diverged, followed by clade A3 containing *P. gracilis* subsp. *nuttallii* from OK and *P. gracilis* subsp. *gracilis* from TX with 100% support (Figs 2.3 and 2.4). Finally, *P. angustifolia* from OK formed a monophyletic group (A4), which was placed as sister to the clade containing two subgroups: A5 and A6 (Figs 2.3 and 2.4).

Western clade – The western data set that included putative polyploids comprised 84 samples representing 46 species. The total number of loci in the assembly was 38,212 loci, of which 39,016 were variant sites (because some loci had more than one variant site). The number of loci in the assembly varied from 708 to 11,468 with an average of 4,182 loci per sample, such that the proportion of missing data per sample ranged from 69.98% to 97.28% with an average of 88.96%. We removed samples with high missing data and rooted the tree with *Paysonia lasiocarpa* subsp. *berlandieri*.

Overall, the base of the phylogeny of the western clade was made up of a grade of 13 single accessions and small clades that were placed as successive sisters, many with strong support, to the remainder of the accessions. In general, accessions at the base of the tree were sampled from southern localities such as Mexico and Texas. These clades and accessions, in order of divergence, were made up of 1) *P. wyndii* from Mexico placed as the sister group to a strongly supported clade (100 % BS) containing the rest of the species in the Western Clade, followed by 2) one accession of *P. argyraea* subsp. *diffusa* from Mexico, 3) a strongly supported clade A1 (97% BS) containing two *P. ovalifolia subsp. ovalifolia* accessions from NM and OK, 4) one accession of *P. gordonii* from NM, 5) another accession of *P. gordonii*, 6) a moderately supported (83% BS) clade A2 containing *P. arizonica* from Arizona (AZ)+ *P. goodingii* from NM (with 100% BS) and *P. gordonii* from OK + *P. ovalifolia* subsp. *ovalifolia* from NM (with 97% BS), 7) Clade A3 contains *P. nelsonii* from Wyoming (WY) + *P. reediana* from Nebraska (NE) with 100% BS support, 8) clade A4 (100% BS) placed *P. crassistigma* from Argentina as the sister group to *P. congesta* from Colorado (CO) + *P. curvipes* from WY (94% BS), 9) in clade A5, *P. montana* and *P. navajoensis*, both from NM placed as strongly supported sisters to *P. brassicoides* accessions from WY with 100% BS, respectively, 10) strongly supported clade A6 (98 % BS) composed of *P. reediana* and one sample of *P. spathulata* from WY placed as successive sisters to *P. pachyphylla* from Montana + *P. spathulata* from WY (with 99% BS), and 11) one accession of *P. rectipes* from NM, which is sister to a moderately supported Clade B (82% BS) containing the remainder of the accessions (Fig. 2.5).

Clade B had several weakly supported branches at the base; the first clade B1 was a weakly supported clade (47% BS) containing *P. vitulifera + P. wardii* (58% BS) placed as sister to *P. purpurea* + *P. wardii* (97% BS) and *P. newberryi* subsp. *newberryi*+ *P. chambersii* +*P.*

navajoensis (with 99% BS), followed by *P. floribunda* subsp. *osterhoutii* placed as a weakly supported (47% BS) sister to a weakly supported clade (54%) containing the rest of the accessions, which was divided into two larger clades (WC1 and WC2) (Fig. 2.5).

The first larger clade (WC1) was weakly supported (67% BS) and contained accessions solely from Utah and Colorado (Fig. 2.4). It was divided into two subclades: the first clade B2 was strongly supported (100% BS) and contained two strongly supported sister groups (100% BS), one containing two accessions of *P. parvula* and the other containing two accessions of *P. pulvinata,* while the second B3 placed *P. navajoensis* as a weakly supported (67% BS) sister group to a strongly supported clade (100% BS) containing three *P. calcicola* species from Colorado and a weakly supported clade that placed *P. intermedia* and *P. hitchcockii* subsp. *rubicundala* as sister groups to a strongly supported clade (91%BS) containing *P. hitchcockii* subsp. *tumolosa* + *P. intermedia,* the latter three were collected from Utah (Fig. 2.5).

The other large clade (WC2) placed one accession of *P. nelsonii* as a weakly supported sister (94% BS) to two large clades (Fig. 2.4). Clade B4 is the first clade and is weakly supported (77% BS) and is split into two weakly supported subclades, one containing another accession of *P. nelsonii* placed as sister to *P. arenosa* + *P. occidentalis* subsp. *occidentalis*, and one in which *P. wardii* and *P. humilis* are placed as strongly supported sisters to *P. eriocarpa* + *P. floribunda* subsp. *floribunda* (Fig. 2.5).

The second large clade in WC2 is strongly supported (98% BS) and is split into two groups B5 and B6 (Fig. 2.5). The first group B5 places *P. grahamii* from Utah as the sister group to two clades, one well supported (98% BS) comprised of *P. kingii* subsp. *kingii* from Oregon + *P. klausii* from Montana, and the second poorly supported (57% BS) and comprised of *P.*

occidentalis subsp. *occidentalis* from Oregon placed as sister to a well-supported clade (100% BS) made up of *P. carinata* subsp. *carinata* from Montana + *P. reediana* (Fig. 2.4).

The second subclade B6 placed two groups, *P. fremontii* from Wyoming + *P. kaibabensis* from Arizona (97% BS) and *P. kingii* subsp. *diversifolia* placed as sister (100% BS) to a strongly supported clade (100% BS) comprised of *P. iveyana* + *P. pinetorum* from New Mexico, as successive sisters two larger clades (Fig. 2.5).

The first larger clade B6(i) placed *P. prostrata* from Idaho, *P. geyeri* subsp. *purpurea* from Idaho and *P. pruinosa* from New Mexico as strongly supported successive sisters to two clades, both of which are strongly supported, one comprised of *P. prostrata* from Utah + *P. pruinosa* from Colorado, and the other comprised of *P. carinata* subsp. *pulchella* from Idaho placed as sister to a strongly supported clade (94% BS) made up of *P. pruinosa* from Colorado + *P. prostrata* from Idaho (Fig. 2.5).

The second larger clade B6 (ii) is well supported (91% BS) and is divided into two groups (Fig. 2.4). The first is well supported and contains two strongly supported subclades (both 100% BS) comprised of *P. subumbellata* from Utah+ *P. vicina* from Colorado, and two *P. valida* from New Mexico (Fig. 2.4). The second places: 1) a clade of *P. lesicii* from Montana + an unknown subspecies of *P. occidentalis* from Idaho, 2) and single accession of *P. occidentalis* subsp. *occidentalis, P. occidentalis subsp. occidentalis, P. kingii* subsp. *latifolia* from Utah, and *P. hitchcockii* subsp. *confluence* from Nevada as successive sister species to the monophyletic *P. carinata* subsp. *pulchella* clade (Fig. 2.5). Overall, most of the species represented by multiple accessions were not monophyletic except *P. brassicoidies, P. calcicola, P. parvula, P. pulvinata, P. valida,* and *P. carinata subsp. pulchella*. Furthermore, out of the 24 *Physaria* sensu stricto

species, 8 species are in the phylogeny and none of these species grouped together except *P. newberryi* subsp. *newberryi* from New Mexico and P. *chambersii* from Utah.

Western clade without potential polyploid species - In this dataset, we removed potential polyploid species from clade 2, reducing it to 66 samples representing 39 species. We removed accessions of *P. argyraea* subsp. *diffusa, P. ovalifolia* subsp. *ovalifolia, P. ovalifolia, P. gordonii, P. vitulifera, P. newberryi* subsp. *newberryi, P. floribunda* subsp. *osterhoutii, P. calcicola, P. arenosa, P. floribunda* subsp. *floribunda, P. kingii* subsp. *latifolia* as the potential polyploid species. The resulting assembly had 69,760 loci, of which 69,299 were variant sites. The number of loci in the assembly varied from 883 to 22,410 loci, with an average of 6,682 loci. The resulting proportion of missing data ranged from 67.87% - 98.73% with an average of 90.42% missing data.

Removal of the polyploid species completely changed the topology of the tree (Figs. 2.6 and 2.7). It reduced the number of taxa in the grade of the base of the tree, such that only *P. wyndii* from Mexico, *P. curvipes* from Wyoming, *P. goodingii* from New Mexico, and *P. crassistigma* from Argentina came as early diverging successive sisters to the rest of the clade, which was strongly supported (100% BS) and divided into two strongly supported groups (Group A and Group B) (Fig. 2.6). The first, with 94% BS, was made up of two clades, in one (clade A1) one accession of *P. pachyphylla* from Montana was nested within a clade made up of two accessions of *P. spathulata* from Wyoming. The second clade A2 was a strongly supported clade (100% BS) that contained two *P. brassicoides* accessions from Wyoming that again formed a monophyletic group with 100% support, which was sister to a clade in which *P. nelsonii* from

Wyoming was nested within a strongly supported clade made up of two accessions of *P. reediana* from Wyoming and Nebraska (99% BS; Fig. 2.6).

In the second clade, a group (A3) made up of *P. carinata subsp. carinata, P. klausii* from Montana, and *P. occidentalis* subsp. *occidentalis* from Oregon followed by a single accession of *P rectipes* were placed as strongly supported successive sisters to the remainder of the clade (Clade B), which was divided into two major clades (Fig. 2.6).

The first major clade (B1) showed three pairs of taxa were all placed as successive sisters to two subclades (B2 and B3), subclade B2 was divided into two groups including *P. hitchcockii* subsp. *tumulosa* from Utah + *P. montana* from NM, and *P. hitchcockii* subsp. *rubicundula* from Utah as sister to a monophyletic group two of *P. intermedia* accessions (Fig. 2.6). Subclade B3 placed *P. navajoensis* as sister to a clade containing two groups, one weakly supported (77% BS) comprising *P. navajoensis* + *P. newberryi* subsp. *newberryi* from NM, and the other placing P. *chambersii* + *P. navajoensis* from NM as sister to a clade in which *P. purpurea* from AZ was nested within two accessions of *P. wardii* from AZ, which is sister to a clade containing *P. purpurea* + *P. wardii.* Therefore both *P. navajoensis* and *P. wardii* are paraphyletic group (Fig. 2.6).

In the second major clade (WC2), clade B4 made up of *P. eriocarpa* from Montana + *P. humilis* from Montana and *P. reediana* + *P. nelsonii* from WY + *P. occidentalis* subsp. *occidentalis* from Idaho (ID) was placed as sister to the remainder of the clade. Next, a grade of small clades, B5 through B10, including 1) *P. grahamii* + *P. prostrata*, both from Utah (B5), 2) *P. wardii* from Utah placed as sister to *P. fremontii* from WY + *P. kaibabensis* from AZ (B6), 3) *P. pinetorum* + *P. iveyana,* both from NM (B7), 4) a clade with 100% BS comprised of *P. subumbellata* from Utah+ *P. vicina* from CO (B8), and a monophyletic group of two *P. vallida* from CO (B9), 5)

two *P. occidentalis* from ID formed a clade (B10) with 97% support, and 6) *P. lesicii* from Montana were placed as successive sister to the rest of the species in the clade, which divided into two small clades (B10 and B11) (Fig. 2.6). The first clade (B11) which was weakly supported (79% BS) was comprised of *P. occidentalis* subsp. *occidentalis*, and *P. hitchcockii* subsp. *confluens*, placed as successive sisters to a strongly supported monophyletic group comprised of two accessions of *P. carinata* subsp. *pulchella*. The other clade (B12) was moderately supported (84% BS) and comprised of *P. prostrata* from Idaho and *P. geyeri subsp. purpurea* from Idaho, *P. carinata* and *P. pruinosa* from CO and NM (Fig. 2.6). In that clade, *P. prostrata* from Idaho and *P. geyeri subsp. purpurea* from Idaho were placed as successive sisters with strong support (100% BS) to a strongly supported (100%) BS) clade containing *P. carinata* from + *P. prostrata*, both from Idaho, and a strongly supported clade (100%BS) composed of three individuals of *P. pruinosa* from New Mexico and Colorado (Fig. 2.6). In this tree, all of the species that were previously monophyletic continued to be monophyletic when polyploids were removed, plus *P. intermedia* and *P. pruinosa* became monophyletic. In addition, several species became paraphyletic, including *P. spathulata, P. reediana, P. navajoensis*, and *P. wardii*.

2.5. Discussion

In this study, we reconstructed the phylogeny of genus *Physaria* (Brassicaceae) using a 2b-RAD-seq DNA sequencing approach. The current study represents the most comprehensive taxonomic sampling so far for the genus, which included 86 species out of ~108 species described in *Physaria* so far. The goals of current study were to reconstruct a well resolved phylogeny of genus *Physaria* to test whether the species relationships proposed by early taxonomists Payson and Rollins corresponded to the phylogeny, to investigate the monophyly of

species to aid in species delimitation, and to identify and remove the potential polyploid taxa from the phylogeny to investigate how they affect the topology of the phylogeny.

The species tree, which included one sample per species, resulted in two clades with strong support values; the first monophyletic group is a small clade which contained species predominantly from eastern North America, and the second monophyletic clade is a larger clade that contained species that are predominantly distributed in Western North America (Figs. 2.1; 2.8. and 2.9). In eastern clade, *P. argyraea* from Texas and *P. densiflora* from Mexico came at the base of the tree (Fig. 2.8). In the Western clade, *P. wyndii*, *P. argyraea subsp. diffusa* from Mexico came at the base of the tree (Fig. 2.9). This result suggests that the genus potentially originated in southern North America, as predicted by Payson. However, assignment to the two major clades and the biogeographic patterns that show species in Mexico and Texas at the base of these clades both rely heavily on the rooting of tree with a single outgroup, which sequenced relatively poorly. Thus, including additional outgroups is needed. For the sub-analyses, one of the species from the other clade could also be used as the outgroup once the root of the tree is clarified.

Within the eastern and western clades, the resulting phylogeny often grouped species collected from nearby locations into clades irrespective of their taxonomic placement (Figs 2.8 and 2.9), therefore we can see a strong biogeographical affinity towards species grouping, possibly due to hybridization within geographic regions. Because the molecular phylogeny is not congruent with the traditional morphology-based phylogeny, it demonstrates that the traditional morphological characters that used to delimit species may exhibit some homoplasy (Bridge et al., 2023). Although some of the species' relationships are like those found in previous studies, most of the

species included into the current phylogeny were not included in previous analyses, such that the results present a novel picture of the evolutionary relationships in the group.

2.5.1 Species relationships proposed by Payson 1921 and Rollins and Shaw 1973

In his 1921 monograph, *Payson* grouped species into the Sections *Alyssum*, *Enantiocarpa* and *Eulesquerella* with one, three and forty-eight species, but only species in section *Eulesquerella* are currently placed in the genus *Physaria*. Payson also proposed many informal species groups (Table 2.1, Fig. 2.10). In the phylogeny generated in the present study, two of the three species in the group are placed in the eastern clade, but none of the other species in the informal groups were placed together in the phylogeny, rejecting the hypothesis that species would be grouped according to the groups proposed by Payson (Fig. 2.11). Payson relied on morphological features of the pod, pedicel and leaves, number of ovules and the branching pattern of stellate trichomes to form these groups but basing these grouping on just a few morphological characters might not be enough for inferring species relationships. However, our results do support Payson's biogeographic hypothesis that the genus originated in Mexico and Texas.

Subsequently, Rollins and Shaw grouped *Physaria* species into 10 informal groups that differed from those proposed by Payson (Table 2.1). Species in group 1 were almost all placed in the eastern clade 1 except *P. gordonii*, which placed in the western clade. All the species in group 7 are placed in the western clade except *P. globosa* (Fig. 2.12). Out of the six species in group 5, we included four species in the phylogeny and three of them were placed in the western clade except *P. inflata*, which is placed in the eastern clade (Fig. 2.12). Out of the six species in group 4, we included four species in the present study and three of them are placed in the western clade, except *P. engelmannii* was placed in the eastern clade (Fig. 2.12). Two species that were

included in group six were both placed in the western clade (Fig. 2.12). However, the assignment of species to the eastern clade and western clade is based on rooting the tree with a single outgroup from Mexico, which sequenced relatively poorly. It will be necessary to include additional outgroups to confidently root the tree and to verify the placement of taxa into major clades.

2.5.2. Monophyletic species in genus *Physaria*

Our second objective was to assess the monophyly of the species represented by multiple accessions in the phylogeny. Most of the species in the eastern clade were monophyletic except a few species, several of which are putative polyploid species. Both when the phylogeny included and excluded putative polyploid species, *P. recurvata*, *P. gracilis*, *P. angustifolia*, *P. globosa,* and *P. ouachitensis* in the eastern clade each formed monophyletic groups with 100% bootstrap support value. Furthermore, *P. filiformis* and *P. globosa* were placed as sister species to each other, a relationship that was also supported by previous studies (Fuentes Soriano and Kellogg, 2021). However, relationships at the base of the clade differed when polyploids were removed, possibly because some polyploids were of hybrid origin, such as *P. argyraea*.

In the western clade, when the phylogeny included potential polyploid species, some species formed monophyletic groups, including 1) *P. brassicoides,* 2) *P. parvula* from Colorado and Utah 3) *P. pulvinata* from Colorado, 4) all three accessions of *P. calcicola* form Colorado, 5) two *P. valida* accessions from New Mexico, 6) and two *P. carinata* subsp. *pulchella* accessions from Idaho and New Mexico. In contrast, several species, such as *P. nelsonii* and *P. occidentalis* were placed as sister to different species in the phylogeny forming polyphyletic groups. Without potential polyploid species and potential hybrid species included, the phylogeny of the western

clade resulted in the same monophyletic species as the phylogeny that included putative polyploid species. Additionally, *P. pruinosa* and *P. intermedia* formed a monophyletic group when polyploids were removed, suggesting that they could be potential progenitors of polyploids. Several species became paraphyletic when polyploids were removed, including *P. spathulata, P. reediana, P. navajoensis*, and *P. wardii*; additional taxonomic work is necessary to evaluate whether these taxa may need re-circumscription.

Overall, many of the species' relationships in our phylogeny did not correspond to previous morphological circumscription or to previous phylogenies of genus *Physaria*. One difference between the previous and present phylogenetic studies is that previous studies did not include as many taxa as the present study. However, given the strong geographical signature of species relationships in the study, the results could have been the produce of hybridization within geographical regions. Alternatively, another possibility is that species have experienced convergent morphological evolution (Davalos et al., 2012), leading to the similar morphologies being present in multiple geographic locations, which could have resulted in different species with the same morphology (Zou and Zhang 2016); additional research is needed to test the merits of these hypotheses. As many species likely diversified within a short period of time, this rapid accumulation of species could be due to many ecological and evolutionary processes such as geographic isolation, sexual selection, or adaptive radiation. Rapidly diversifying lineages possess significant challenges in phylogeny reconstruction including incomplete lineage sorting which limits the phylogenetic signal (Bagley et al 2020).

2.5.3. Change in the topology of the eastern clade and western clade depending on the inclusion or exclusion of the polyploid species.

In both clades, the removal of polyploid taxa affected the topology. Although the monophyletic species in the eastern clade remained the same when polyploid species were removed, the

topology at the base of the eastern phylogeny changed (Fig. 2.4); for example, in the phylogeny with polyploids, one accession of *P. sessilis* was placed as sister to the subclade of *P. filiformis* + *P. gracilis*, *P. angustifolia*, and *P. ouachitensis* with high support (90%), whereas in the phylogeny without polyploid species, *P. sessilis* was placed as sister to *P. recurvata* with low support (73%). It also completely changed many aspects of the topology of the western clade; for example, removal of the polyploids reduced the grade at the base of the phylogeny (Fig. 2.7). Often, hybrids are placed at the base of a tree, suggesting that some of the polyploids that were removed may be of hybrid origin.

Additional research is needed to understand how polyploids have affected the phylogenies in this study. For example, because we removed all known putative polyploid taxa all at once, it is unclear which species had the greatest effect on the topology. Additional analysis involving removal of one polyploid at a time may help clarify which species affect the topology most strongly, to help identify which species could be of potential hybrid origin. Furthermore, another issue with our analyses is that we inferred the ploidy level of *Physaria* species through the available literature, but the majority of the species do not have ploidy counts, and we did not conduct ploidy estimates of the samples included in the present study, such that the ploidy of many samples included in the study are unknown. Therefore, the use of molecular methods to estimate ploidy or additional chromosome counts are needed to measure the ploidy level of the species included in this study, which would greatly facilitate phylogenetic analysis and the assessment of its impact on the tree building. These counts would also facilitate an analysis of the evolution of chromosome numbers in *Physaria*, which would be helpful for understanding how variation in chromosome number and polyploidy have affected the evolutionary history of the group.

Polyploidy may also be partially responsible for another prominent feature of our phylogenies, which is the non-monophyly of most species. Given that species are generally morphologically cohesive, the incongruence between the phylogeny and the morphology-based species boundaries could also be because we have included previously unrecognized hybrids and allopolyploids in the phylogeny, which may disrupt species monophyly (Sancho et al., 2022). However, the non-monophyly of species could also be due to horizontal gene transfer and incomplete lineage sorting (Cellinese et al., 2012). Methodical reasons for conflict between morphology and molecular data may also be due to a high proportion of missing data in the sequences; therefore, additional sequencing needs to be conducted to overcome those problems. The software we used to call SNPs is designed to assess the genotype of diploid organisms, therefore the higher level of heterozygosity of polyploids could be masked and they have been considered diploids with two alleles. Finally, inclusion of successfully sequenced outgroups from several genera of tribe Physariae would be another step towards improving the current phylogeny. Additional research is needed to test the relative merits of these hypotheses in affecting species relationships and species monophyly.

2.6. Conclusions

The current study reconstructed a well resolved phylogeny including many taxa of genus *Physaria* to date. The resulting phylogeny revealed two main clades, mainly placing species based on their geographical location. The grouping of the species follows a strong biogeographical pattern irrespective of their taxonomic placements. This confusing placement of some species may be due to hybridization, incomplete lineage sorting, or other factors, which needs further investigation with more field sampling. The study incorporated the ploidy level of *Physaria* species, and removal of putative polyploids resulted in major changes in the topology

or the support values in the branches in the phylogeny, suggesting that some polyploids are likely of hybrid origin; however, additional study on the origins of each polyploid species is needed, as some of them represent species complexes that will require more extensive population genomic analyses to elucidate. Further studies are also needed to assess the ploidy level of additional species to determine whether any of the species in the phylogeny are of previously unrecognized polyploid origin. Based on the phylogeny, we identified taxa that need more sampling to increase the resolution of the phylogeny and to identify more monophyletic groups, as well as some taxa that did not sequence well that will need further sequencing attempts. However, the present study provides an important phylogenetic framework that can serve as the basis for future studies on the biogeography of genus *Physaria,* which has an interesting geographical distribution and includes many rare and endangered species.

REFERENCES

Acha, S., Linan, A., MacDougal, J., & Edwards, C. (2021). The evolutionary history of vines in Neotropical biodiversity hotspot: Phylogenomics and biogeography of a large passionflower clade (*Passiflora* section Decaloba). *Molecular Phylogenetics and Evolution*, *164*, 107260.

Al-Shehbaz, I. A., & O'Kane, S. L. (2002). Lesquerella Is United with *Physaria* (Brassicaceae). *Novon, 12*(3), 319. https://doi.org/10.2307/3393073

Bagley, J. C., Uribe-Convers, S., Carlsen, M. M., & Muchhala, N. (2020). Utility of targeted sequence capture for phylogenomics in rapid, recent angiosperm radiations: Neotropical *Burmeistera* Bellflowers as a case study. *Molecular Phylogenetics and Evolution, 152*, 106769. https://doi.org/10.1016/j.ympev.2020.106769

Beardsley, P. M., Schoenig, S. E., Whittall, J. B., & Olmstead, R. G. (2004). Patterns of evolution in western North American *Mimulus* (Phrymaceae). *American Journal of Botany, 91*(3), 474–489. https://doi.org/10.3732/ajb.91.3.474

Bridge, T. C., Cowman, P. F., Quattrini, A. M., Bonito, V. E., Sinniger, F., Harii, S., Head, C. E., Hung, J. Y., Halafihi, T., Rongo, T., & Baird, A. H. (2023). A tenuis relationship: Traditional taxonomy obscures systematics and biogeography of the 'Acropora tenuis' (scleractinia: Acroporidae) species complex. *Zoological Journal of the Linnean Society*.

Cellinese, N., Baum, D. A., & Mishler, B. D. (2012). Species and Phylogenetic Nomenclature. *Systematic Biology, 61*(5), 885–891. https://doi.org/10.1093/sysbio/sys035

Dávalos, L. M., Cirranello, A. L., Geisler, J. H., & Simmons, N. B. (2012). Understanding phylogenetic incongruence: Lessons from phyllostomid bats. *Biological Reviews, 87*(4), 991–1024. https://doi.org/10.1111/j.1469-185X.2012.00240.x

Doyle, J. & Doyle, J. L. (1987). Genomic plant DNA preparation from fresh tissue- CTAB method. *Phytochem Bull, 19*(11), 11-15.

Eaton, D.A.R., Overcast, I. (2020). Ipyrad: Interactive assembly and analysis of RADseq datasets. *Bioinformatics 36*, 2592–2594. https://doi.org/10.1093/bioinformatics/btz966

Fuentes-Soriano, S., & Al-Shehbaz, I. (2013). Phylogenetic Relationships of Mustards with Multiaperturate Pollen (Physarieae, Brassicaceae) based on the Plastid ndhF gene: Implications for Morphological Diversification. *Systematic Botany*, *38*(1), 178-191. doi: 10.1600/036364413x661854

Fuentes-Soriano, S., & Kellogg, E. A. (2021). Molecular Systematics of Tribe Physarieae (Brassicaceae) Based on Nuclear ITS, LUMINIDEPENDENS, and Chloroplast ndhF. *Systematic Botany*, *46*(3), 611–627. https://doi.org/10.1600/036364421X16312067913318

Grady, B. R., & O'Kane, S. L. (2007). New Species and Combinations in Physaria (Brassicaceae) from Western North America. Novon: *A Journal for Botanical Nomenclature, 17*(2), 182. https://doi.org/10.3417/1055-3177(2007)17[182:NSACIP]2.0.CO;2

Hoang, D. T., Chernomor, O., Von Haeseler, A., Minh, B. Q., & Vinh, L. S. (2018). UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Molecular Biology and Evolution, 35*(2), 518–522. https://doi.org/10.1093/molbev/msx281

Kiefer, M., Schmickl, R., German, D. A., Mandáková, T., Lysak, M. A., Al-Shehbaz, I. A., Franzke, A., Mummenhoff, K., Stamatakis, A., & Koch, M. A. (2014). BrassiBase: Introduction to a Novel Knowledge Database on Brassicaceae Evolution. *Plant and Cell Physiology*, *55*(1), e3–e3. https://doi.org/10.1093/pcp/pct158

Linan, A., Schatz, G., Lowry, P., Miller, A., & Edwards, C. (2019). Ebony and the Mascarenes: the evolutionary relationships and biogeography of *Diospyros* (Ebenaceae) in the western Indian Ocean. *Botanical Journal Of The Linnean Society*, *190*(4), 359-373.

Lysak, M. A., Koch, M. A., Beaulieu, J. M., Meister, A., & Leitch, I. J. (2008). The Dynamic Ups and Downs of Genome Size Evolution in Brassicaceae. *Molecular Biology and Evolution*, *26*(1), 85–98. https://doi.org/10.1093/molbev/msn223

Mazie, A. R., & Baum, D. A. (2016). Clade-specific positive selection on a developmental gene: BRANCHLESS TRICHOME and the evolution of stellate trichomes in *Physaria* (Brassicaceae). *Molecular Phylogenetics and Evolution*, 100, 31–40. https://doi.org/10.1016/j.ympev.2016.03.027

Minh, B. Q., Nguyen, M. A. T., & Von Haeseler, A. (2013). Ultrafast Approximation for Phylogenetic Bootstrap. *Molecular Biology and Evolution*, *30*(5), 1188–1195. https://doi.org/10.1093/molbev/mst024

Mulligan, G. A. (1968). *Physaria didymocarpa*, *P. brassicoides,* and *P. floribunda* (Cruciferae) and their close relatives. *Canadian Journal of Botany*, *46*(6), 735–740. https://doi.org/10.1139/b68-101

Nguyen, L.-T., Schmidt, H. A., Von Haeseler, A., & Minh, B. Q. (2015). IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. *Molecular Biology and Evolution*, *32*(1), 268–274. https://doi.org/10.1093/molbev/msu300

O'Kane S. L. *Physaria*. In: Flora of North America Editorial Committee, eds. 1993+. Flora of North America North of Mexico [Online]. 22+ vols. New York and Oxford. Vol.7.

O'Kane, S. L., Al-Shehbaz, I. A., & Turland, N. J. (1999). (1393) Proposal to conserve the name *Lesquerella* against *Physaria* (Cruciferae). *TAXON*, *48(*1), 163–164. https://doi.org/10.2307/1224642

Payson, E.B. (1921). A monograph of the genus *Lesquerella*. *Ann. Missouri Bot. Gard. 8*:104–236.

Queiroz, K., & Donoghue, M. J. (1988). Phylogenetic systematics and the species problem. *Cladistics*, *4*(4), 317–338. https://doi.org/10.1111/j.1096-0031.1988.tb00518.x

Rice et al. 2015. The Chromosome Counts Database (CCDB) – a community resource of plant chromosome numbers. New Phytol. 206(1): 19-26.

Rollins, R. C. (1993). *The Cruciferae of Continental North America: Systematics of the Mustard Family from the Arctic to Panama.* (pp. 589-714). Stanford University Press.

Rollins, R. C. (1939). The cruciferous genus *Physaria*. *Rhodora 41*: 391-414.

Rollins, R. C., & Rüdenberg, L. (1979). Chromosome numbers of Cruciferae. IV. Contributions from the Gray Herbarium of Harvard University

Rollins, R. C. (1981). Studies in the Genus *Physaria* (Cruciferae). *Brittonia 33* (3): 332-341

Rollins, R., & Shaw, E. (1973). *The Genus Lesquerella (Cruciferae) in North America.* (pp. 4-47). Cambridge, MA: Harvard University Press.

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., Von Haeseler, A., & Jermiin, L. S. (2017). ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nature Methods*, *14*(6), 587–589. https://doi.org/10.1038/nmeth.4285

Salywon, A. M., Rebman, J. P., & Dierig, D. A. (2022). Documented chromosome number determinations in some *Physaria* species (Brassicaceae). *Journal of the Botanical Research Institute of Texas, 16*(2), 499–504. https://doi.org/10.17348/jbrit.v16.i2.1262

Sancho, R., Inda, L. A., Díaz-Pérez, A., Des Marais, D. L., Gordon, S., Vogel, J. P., Lusinska, J., Hasterok, R., Contreras-Moreira, B., & Catalán, P. (2022). Tracking the ancestry of known and 'ghost' homeologous subgenomes in model grass *Brachypodium* polyploids. *The Plant Journal*, *109*(6), 1535–1558.

Štorchová, H., Hrdličková, R., Chrtek, J., Tetera, M., Fitze, D., & Fehrer, J. (2000). An improved method of DNA isolation from plants collected in the field and conserved in saturated NaCl/CTAB solution. *TAXON*, *49*(1), 79–84. https://doi.org/10.2307/1223934

Tropicos.org. Missouri Botanical Garden. 20 Sep 2022 https://tropicos.org

Véron, S., Saito, V., Padilla-García, N., Forest, F., & Bertheau, Y. (2019). The Use of Phylogenetic Diversity in Conservation Biology and Community Ecology: A Common Base but Different Approaches. *The Quarterly Review of Biology*, *94*(2), 123–148. https://doi.org/10.1086/703580

Waite, S. B. (1973). A taxonomic revision of *Physaria* (Cruciferae) in Utah. *Great Basin Naturalist*, *33*, 31-36.

Wang, S., Meyer, E., Mckay, J.K., Matz, M. V. (2012). 2b-RAD: A simple and flexible method for genome-wide genotyping. *Nature Methods 9,* 808–810.

Wood, T. E., Takebayashi, N., Barker, M. S., Mayrose, I., Greenspoon, P. B., & Rieseberg, L. H. (2009). The frequency of polyploid speciation in vascular plants. *Proceedings of the National Academy of Sciences*, *106*(33), 13875–13879. https://doi.org/10.1073/pnas.0811575106

Warwick, S. I., & Al-Shehbaz, I. A. (2006). Brassicaceae: Chromosome number index and database on CD-Rom. *Plant Systematics and Evolution*, 259(2–4), 237–248. https://doi.org/10.1007/s00606-006-0421-1

Zou, Z., & Zhang, J. (2016). Morphological and molecular convergences in mammalian phylogenetics. *Nature Communications*, *7*(1), 12758. https://doi.org/10.1038/ncomms12758

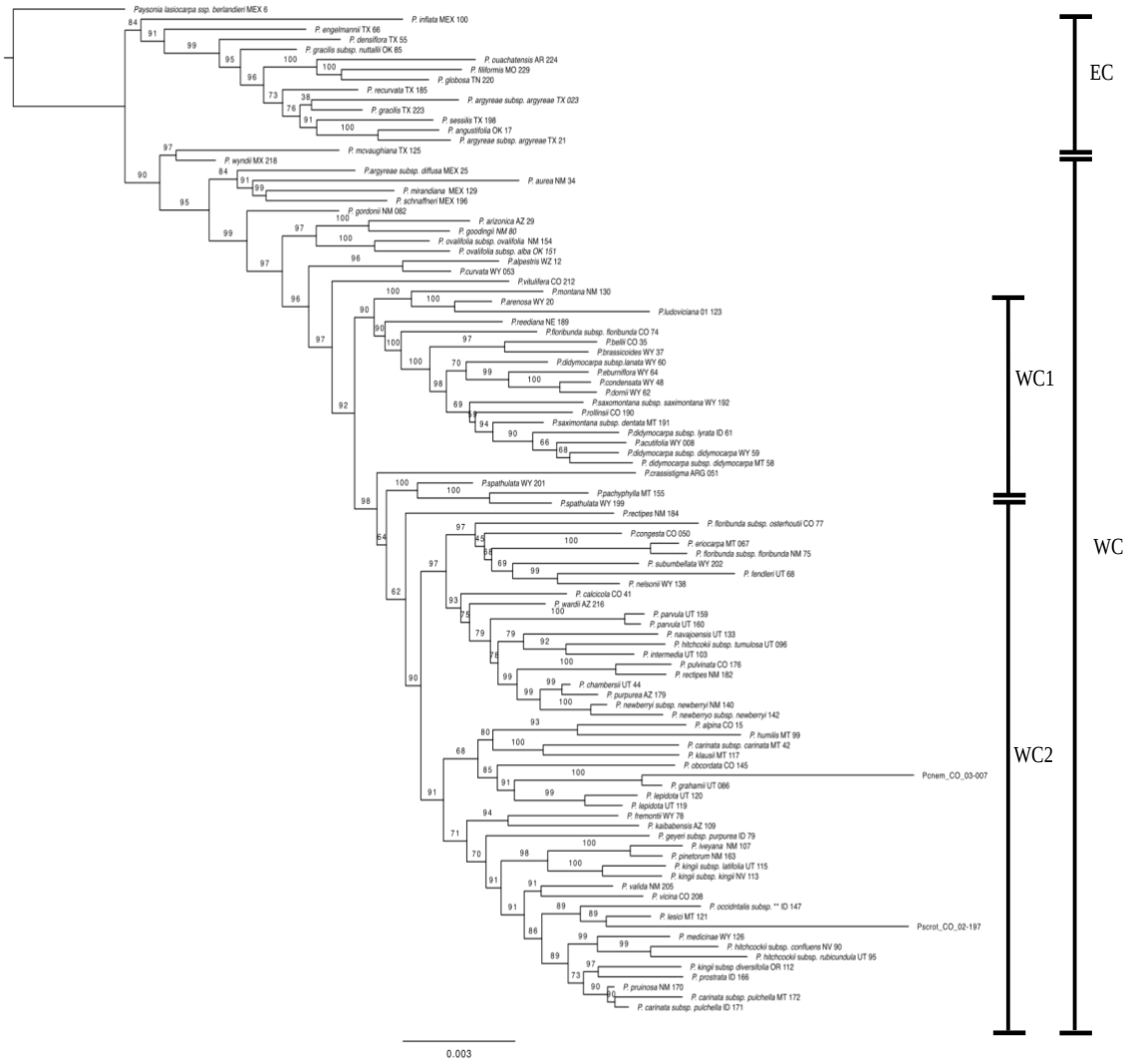

Figure 2.1.Maximum likelihood tree including one sample representing each species of genus Physaria resulted from IQ Tree analysis. Two major clades are annotated. EC – eastern clade, WC- western clade and western clade divided into two subclades as W C1- western clade 1 and WC2 – western clade 2. The numbers above the branches indicate the bootstrap support (BS) values.

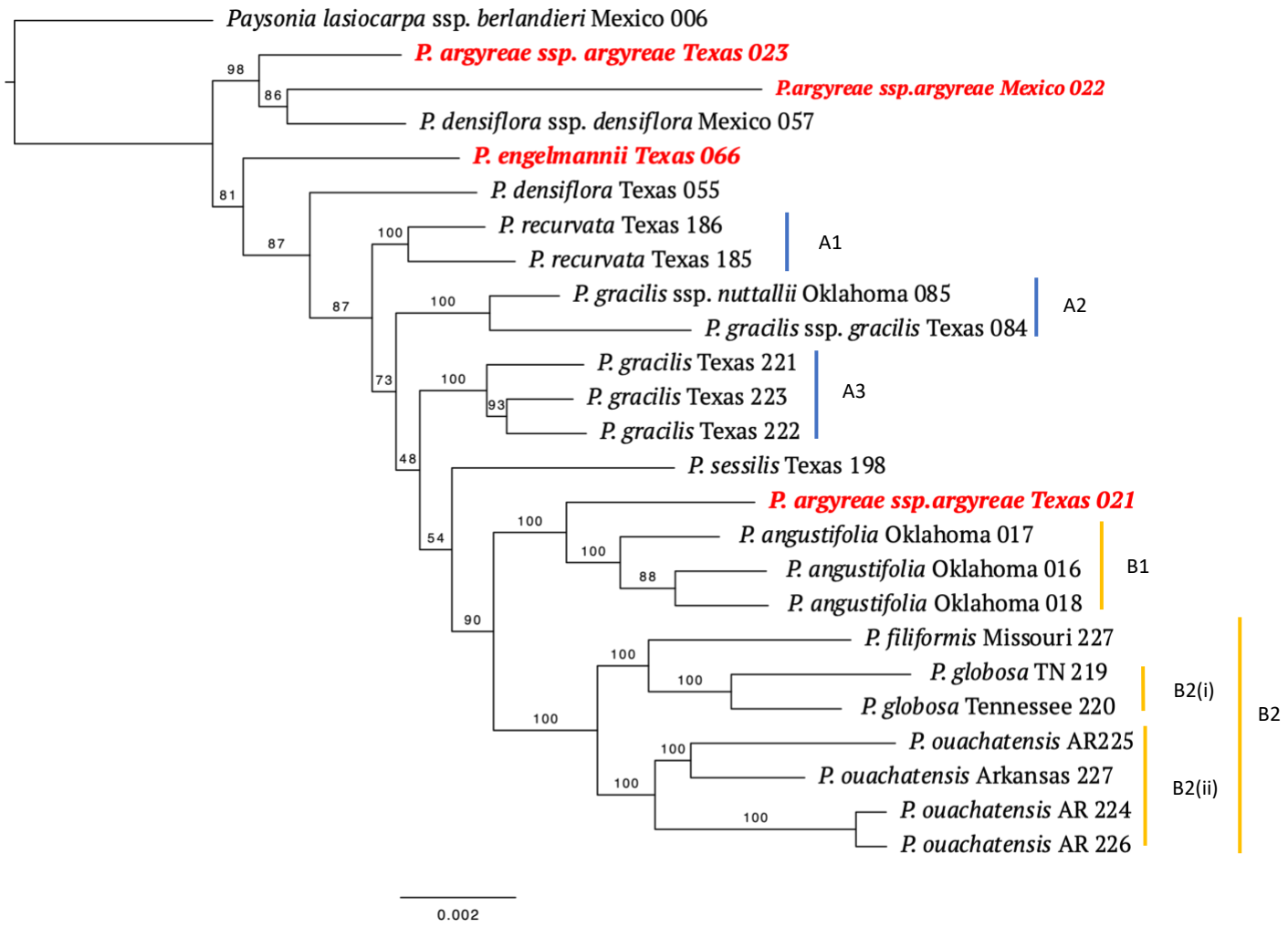

Figure 2.2. Maximum likelihood phylogeny of Eastern clade resulting from IQ tree analysis including potential polyploid species. The tree is rooted with the outgroup *Paysonia lasiocarpa* subsp. *berlandieri*. Bootstrap support values are reported above the branches.

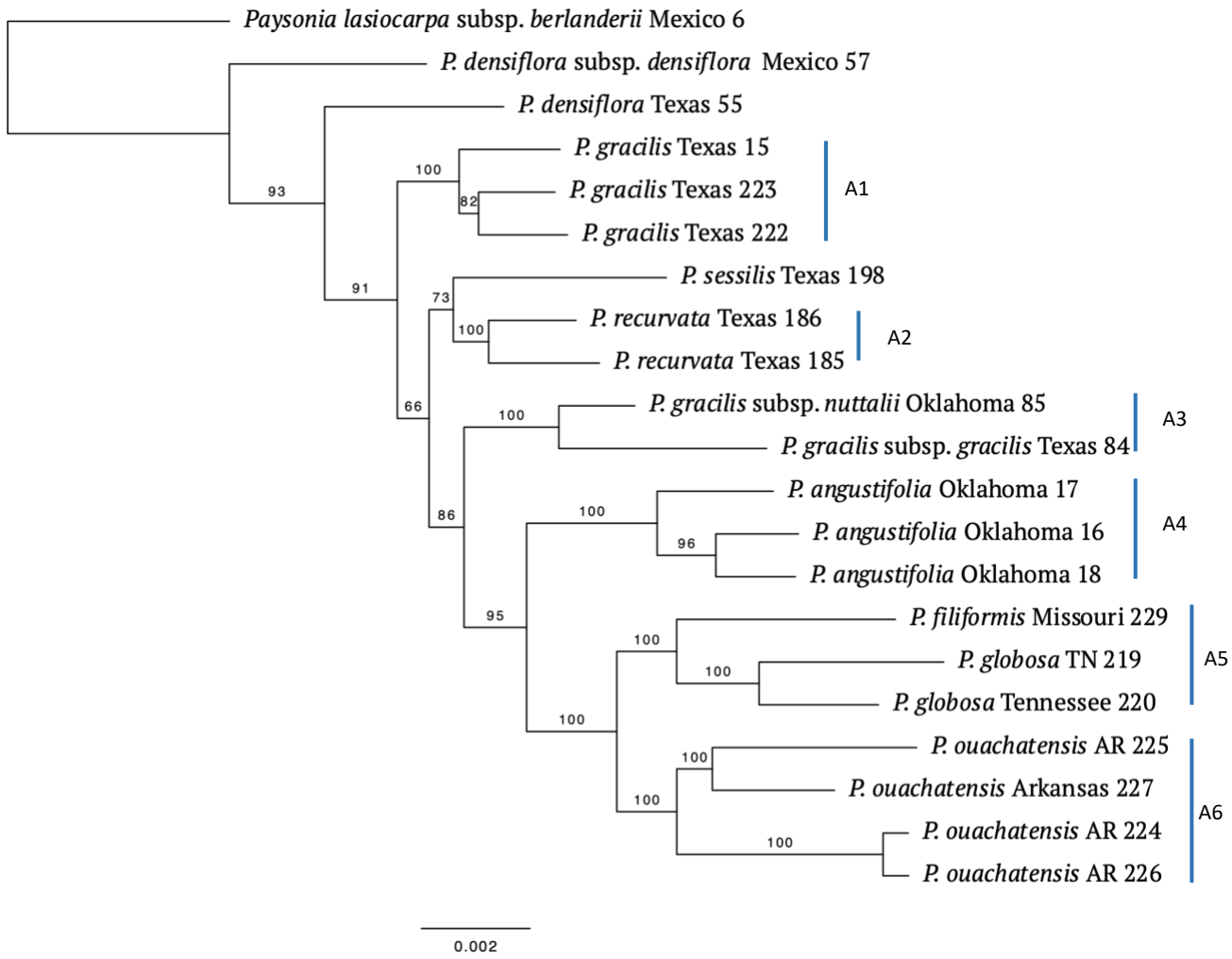

Figure 2.3.Maximum likelihood phylogeny of Eastern clade without potential polyploid species. The support values on the branches are bootstrap support values.

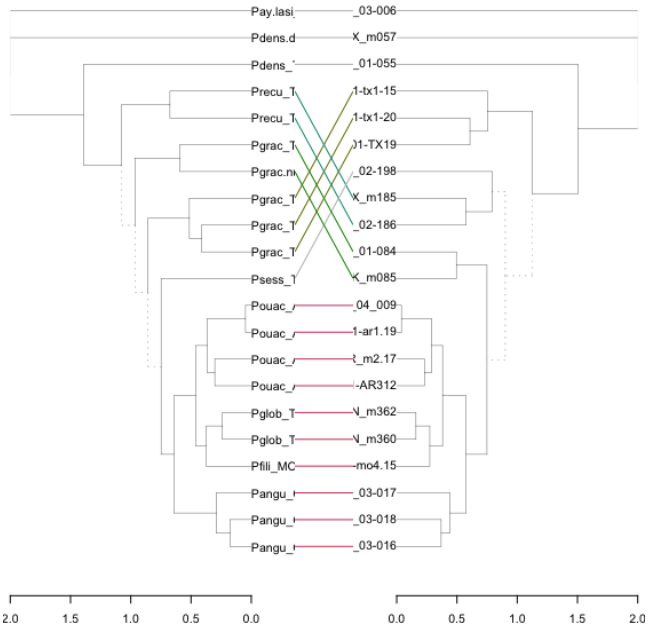

Figure 2.4. Tangelgram comparison of phylogenies of eastern clade from IQ tree without polyploids (left) and with polyploids (right). Taxa that are not present in both trees (i.e.,) polyploids are pruned.

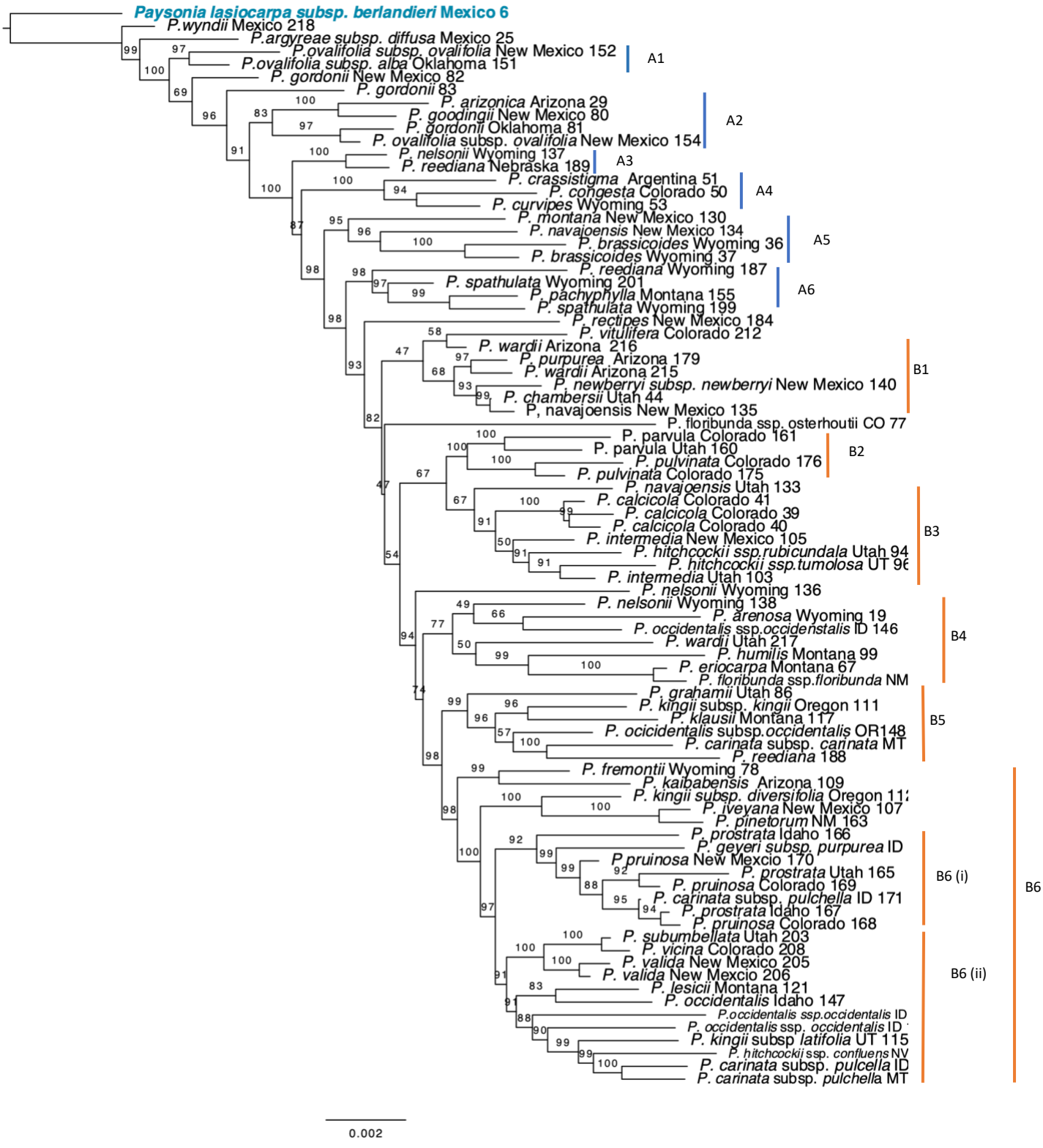
Paysonia lasiocarpa subsp. berlandieri Mexico 6
P.wyndii Mexico 218
P.argyreae subsp. diffusa Mexico 25
P.ovalifolia subsp. ovalifolia New Mexico 152
P.ovalifolia subsp. alba Oklahoma 151
P. gordonii New Mexico 82
P. gordonii 83
P. arizonica Arizona 29
P. goodingii New Mexico 80
P. gordonii Oklahoma 81
P. ovalifolia subsp. ovalifolia New Mexico 154
P. nelsonii Wyoming 137
P. reediana Nebraska 189
P. crassistigma Argentina 51
P. congesta Colorado 50
P. curvipes Wyoming 53
P. montana New Mexico 130
P. navajoensis New Mexico 134
P. brassicoides Wyoming 36
P. brassicoides Wyoming 37
P. reediana Wyoming 187
P. spathulata Wyoming 201
P. pachyphylla Montana 155
P. spathulata Wyoming 199
P. rectipes New Mexico 184
P. vitulifera Colorado 212
P. wardii Arizona 216
P. purpurea Arizona 179
P. wardii Arizona 215
P. newberryi subsp. newberryi New Mexico 140
P. chambersii Utah 44
P, navajoensis New Mexico 135
P. floribunda ssp. osterhoutii CO 77
P. parvula Colorado 161
P. parvula Utah 160
P. pulvinata Colorado 176
P. pulvinata Colorado 175
P. navajoensis Utah 133
P. calcicola Colorado 41
P. calcicola Colorado 39
P. calcicola Colorado 40
P. intermedia New Mexico 105
P. hitchcockii ssp.rubicundala Utah 94
P. hitchcockii ssp.tumolosa UT 96
P. intermedia Utah 103
P. nelsonii Wyoming 136
P. nelsonii Wyoming 138
P. arenosa Wyoming 19
P. occidentalis ssp.occidenstalis ID 146
P. wardii Utah 217
P. humilis Montana 99
P. eriocarpa Montana 67
P. floribunda ssp.floribunda NM
P. grahamii Utah 86
P. kingii subsp. kingii Oregon 111
P. klausii Montana 117
P. ocicidentalis subsp.occidentalis OR148
P. carinata subsp. carinata MT
P. reediana 188
P. fremontii Wyoming 78
P. kaibabensis Arizona 109
P. kingii subsp. diversifolia Oregon 11
P. iveyana New Mexico 107
P. pinetorum NM 163
P. prostrata Idaho 166
P. geyeri subsp. purpurea ID
P.pruinosa New Mexcio 170
P. prostrata Utah 165
P. pruinosa Colorado 169
P. carinata subsp. pulchella ID 171
P. prostrata Idaho 167
P. pruinosa Colorado 168
P. suburnbellata Utah 203
P. vicina Colorado 208
P. valida New Mexico 205
P. valida New Mexcio 206
P. lesicii Montana 121
P. occidentalis Idaho 147
P.occidentalis ssp.occidentalis ID
P. occidentalis ssp. occidentalis ID
P. kingii subsp latifolia UT 115
P. hitchcockii ssp. confluens NV
P. carinata subsp. pulcella ID
P. carinata subsp. pulchella MT
A1
A2
A3
A4
A5
A6
B1
B2
B3
B4
B5
B6 (i)
B6 (ii)
B6
0.002

Figure 2.5. Maximum likelihood (ML) phylogeny of the species included in the Western clade including potential polyploid species. Bootstrap support values (BS) are shown above branches. The tree is rooted with the outgroup Paysonia lasiocarpa subsp. *berlandieri*.

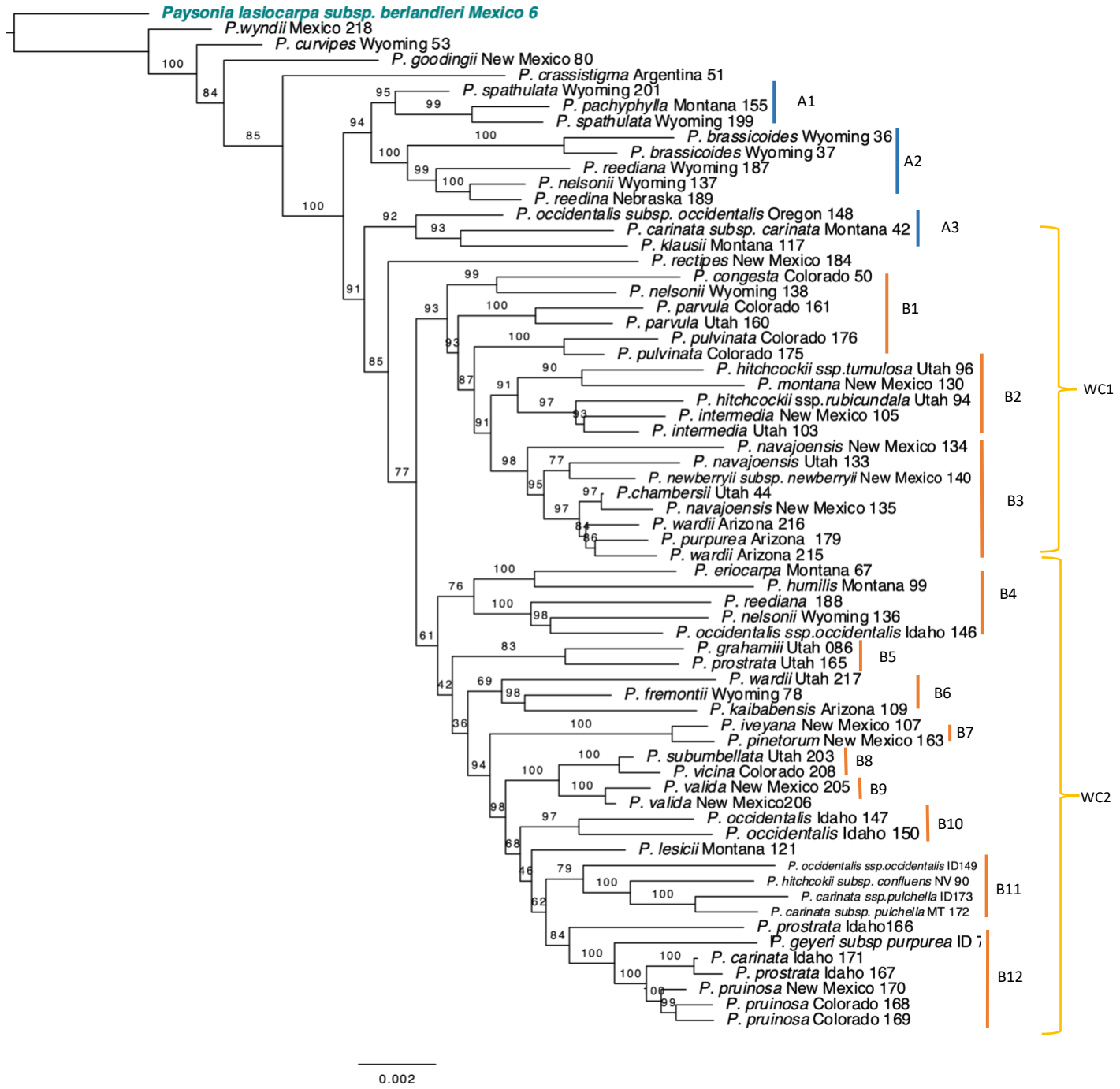

Figure 2.6. Maximum likelihood (ML) phylogeny of the species included in the Western clade without potential polyploids species. Bootstrap support values generated (BS) are shown above branches. The tree is rooted with the outgroup *Paysonia lasiocarpa ssp. berlandieri*

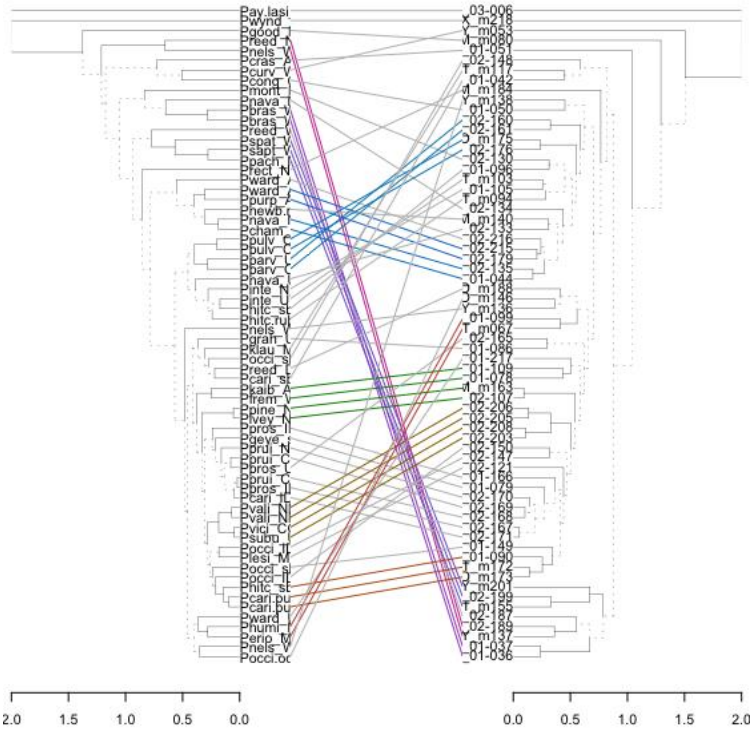

Figure 2.7. Tangelgram comparison of phylogenies of western clade from IQ tree without polyploids (left) and with polyploids (right). Polyploids that are not present in both trees are pruned.

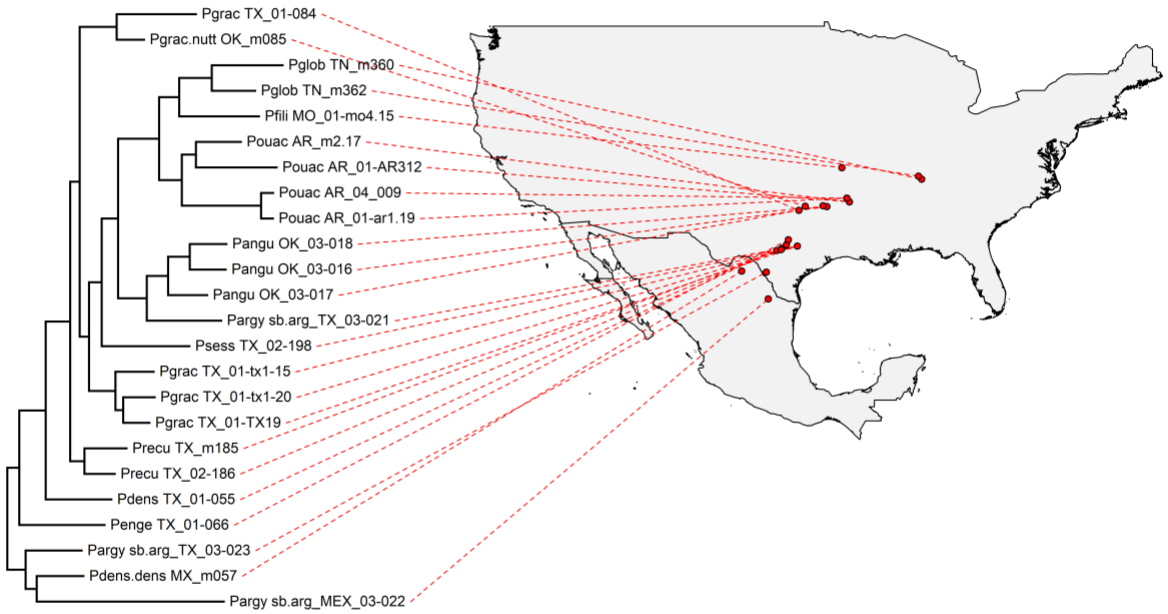

Figure 2.8. Eastern clade phylogeny with polyploid species projected on the geographic map of the sample locations showing the distribution of the species. Species name, collection locality (state), and accession number are provided at the tip of the tree.

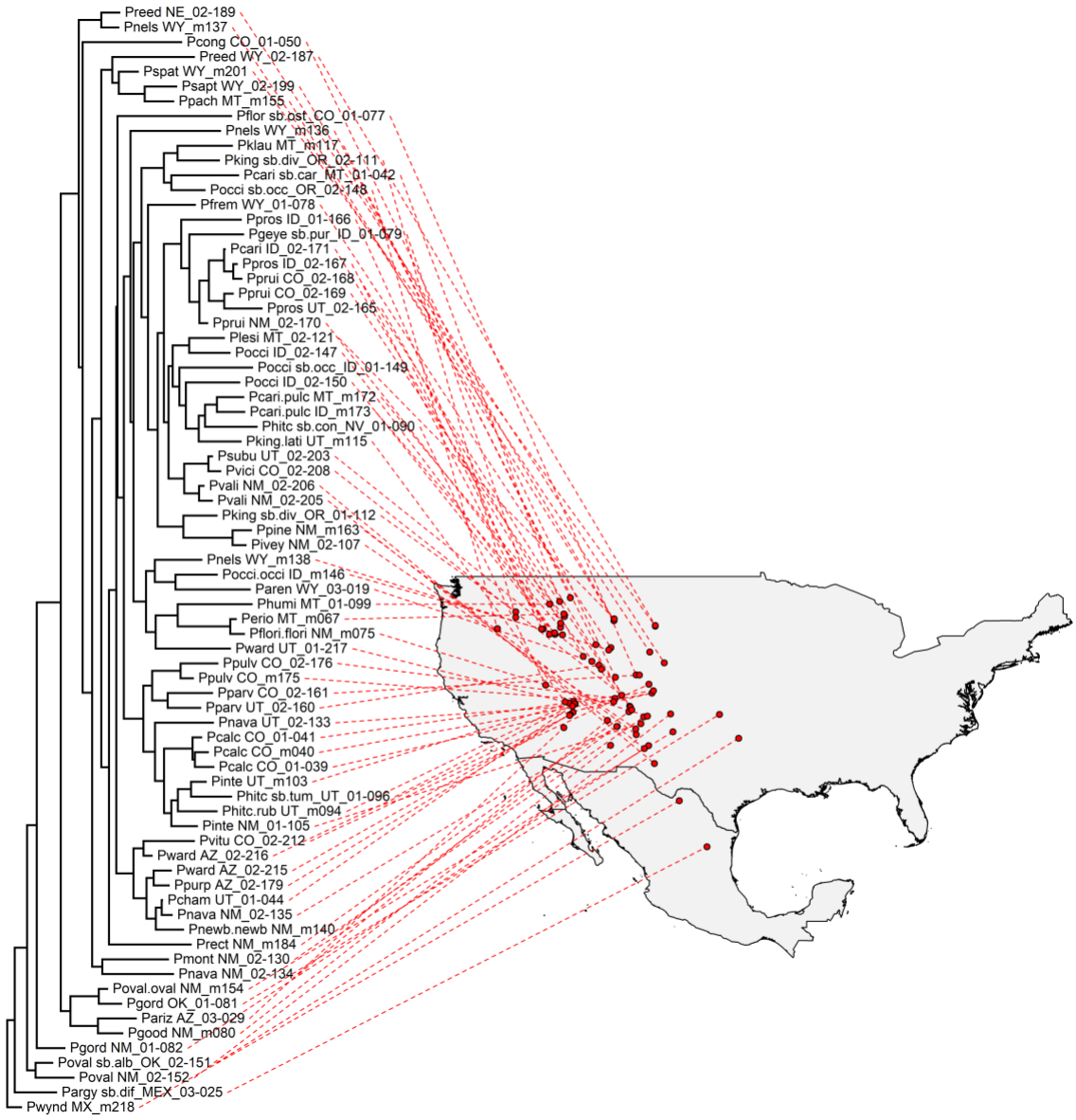

Figure2.9. Western clade phylogeny with polyploid species projected on the geographic map of the sample locations showing the distribution of the species. Species name, collection locality (state), and accession number are provided at the tip of the tree.

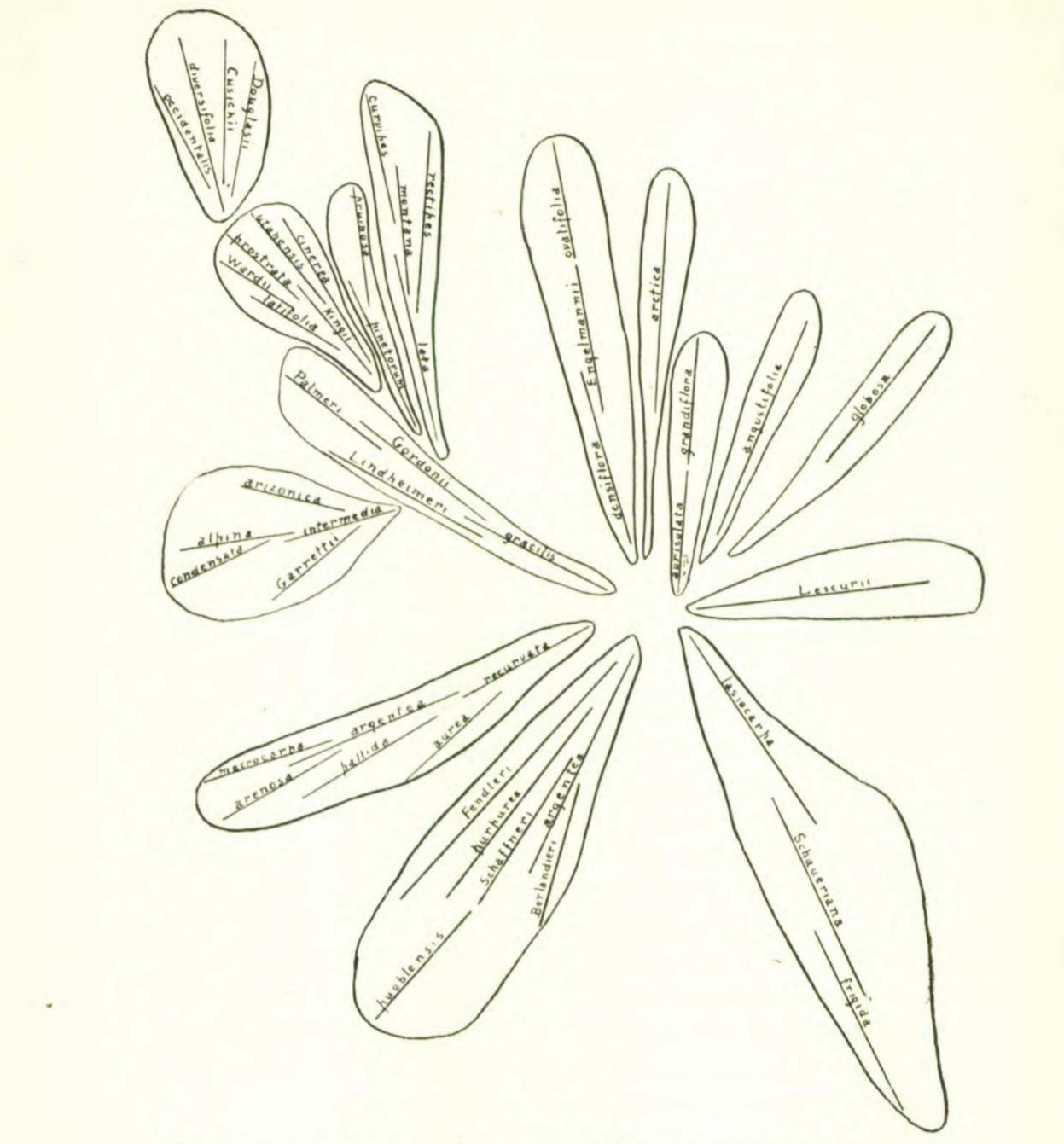

Figure 2.10. Hypothetical species relationships proposed by Payson 1921, in his monograph for the species of genus Lesquerella (later transferred to Physaria). Adapted from Payson 1921.

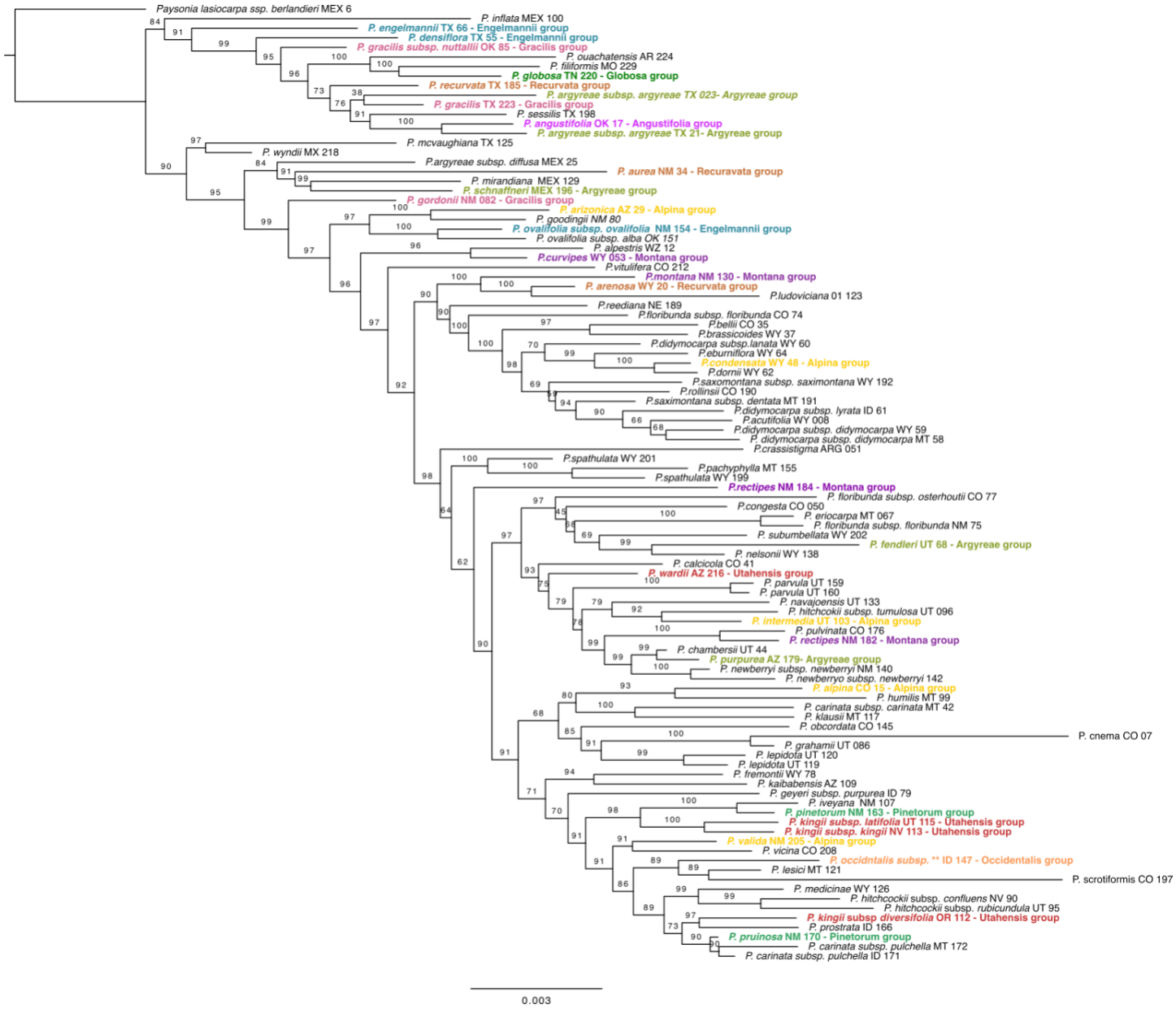

Figure 2.11. Species relationships proposed by Payson in the 1921 monograph are shown on the Maximum likelihood tree including one sample representing each species of genus Physaria.

Figure 2.12. Species relationships proposed by Rollins and Shaw's 1973 monograph are shown on the Maximum likelihood phylogeny including one sample representing each species of genus Physaria.

Table 2.1: Species relationships proposed by Payson 1921 and Rollins and Shaw 1973 for species in genus *Physaria*

Payson 1921: A monograph of genus Lesquerella				Rollins and Shaw 1973: The genus Lesquerella (Cruciferae) of North America		
Sections	Groups	Names of the species	Current status based on the Flora of North America Volume 7 and Tropicos database	Groups	Names of the species	Current status based on the Flora of North America Volume 7 and Tropicos database
Eulesquerella	Engelmannii group	*Physaria densiflora*	Present	Group 1	*Physaria angustifolia*	present
		Physaria engelmannii	Present		*Physaria argyrea*	Present
		Physaria ovalifolia	Present		*Physaria densiflora*	Present
	Arctica group	*Physaria arctica*	Present		*Physaria filiformis*	Present
	Argyraea group	*Physaria argyraea*	Present		*Physaria gordonii*	Present
		Physaria berlandieri	Present		*Physaria gracilis*	Present
		Physaria purpurea	Present		*Physaria lindheimeri*	Present
		Physaria fendleri	Present		*Physaria recurvata*	Present
		Physaria schaffneri	Present		*Physaria sessilis*	Present
		Physaria pueblensis	Present		*Physaria tenella*	Present
	Recurvata group	*Physaria recurvata*	Present		*Physaria thamnophila*	Present
		Physaria pallida	Present	Group 2	*Physaria fendleri*	Present
		Physaria aurea	Present		*Physaria hitchcockii*	Present
		Physaria argentea	present		*Physaria rubicundala*	Present
		Physaria arenosa	Present	Group 3	*Physaria johnstonii*	Present
		Physaria macrocarpa	Present		*Physaria mcvaughiana*	Present
	Angustifolia group	*Physaria angustifolia*	present		*Physaria purpurea*	Present
	Gracilis group	*Physaria gracilis*	Present	Group 4	*Physaria arctica*	Present
		Physaria palmeri	Present		*Physaria calderi*	Present
		Physaria lindheimeri	Present		*Physaria engelmannii*	Present

	Physaria gordonii	Present		*Physaria ovalifolia*	Present
Pinetorum group	*Physaria pinetorum*	Present		*Physaria pinetorum*	Present
	Physaria pruinosa	Present		*Physaria pruinosa*	Present
Montana group	*Physaria lata*	Present	Group 5	*Physaria inflata*	Present
	Physaria rectipes	Present		*Physaria mexicana*	Present
	Physaria montana	Present		*Physaria mirandiana*	Present
	Physaria curvipes	Present		*Physaria pueblensis*	Present
Globosa species	*Physaria globosa*	Present		*Physaria schaffneri*	Present
Mendocina species	*Physaria mendocina*	Present		*Physaria wyndii*	Present
Alpina group	*Physaria intermedia*	Present	Group 6	*Physaria aurea*	Present
	Physaria arizonica	Present		*Physaria gooddingii*	Present
	Physaria alpina	Present	Group 7	*Physaria alpina*	Present
	Physaria condensata	Present		*Physaria arenosa*	Present
	Physaria garrettii	Present		*Physaria arizonica*	Present
	Physaria valida	Present		*Physaria calcicola*	Present
Utahensis group	*Physaria cinerea*	Present		*Physaria cinerea*	Present
	Physaria kingii	Present		*Physaria globosa*	Present
	Physaria wardii	Present		*Physaria intermedia*	Present
	Physaria utahensis	Present		*Physaria lata*	Present
	Physaria prostrata	Present		*Physaria ludoviciana*	Present
Occidentalis group	*Physaria occidentalis*	Present		*Physaria montana*	Present
	Physaria douglasii	Present		*Physaria rectipes*	Present
				Physaria valida	Present
			Group 8	*Physaria carinata*	Present
				Physaria douglasii	Present
				Physaria fremontii	Present
				Physaria garrettii	Present

	Physaria macrocarpa	Present
	Physaria multiceps	Present
	Physaria occidentalis	Present
	Physaria paysonii	present
	Physaria prostrata	Present
	Physaria utahensis	Present
	Physaria wardii	Present
Group 9	*Physaria cordiformis*	Present
	Physaria hemiphysaria	Present
	Physaria kingii	Present
	Physaria palmeri	present
	Physaria peninsularis	present

Table 2.2: Previous estimates of chromosome number and polyploidy for the species in genus *Physaria*

Species names	Sample number	Most common x	Other reported x	Somatic chromosome count (2n)	Diploid (2x)	Triploid (3x)	Tetraploid (4x)	Pentaploid(5x)	Hexaploid (6x)	Reference
Paysonia grandiflora	5	9		18						Brassibase; Rollins 1973
Paysonia lasiocarpa subsp berlandieri	6			14						FNA
Physaria "cnema"	7					*No data available*				
Physaria acutifolia	8	4,5			8,10		16		24	Salywon *et al*., 2022
Physaria alpestris	12			49-52, 52, 67-70 [2n= 48, 64?]						Brassibase; Rollins 1993
Physaria alpina	15									Rollins and Shaw 1973
Physaria angustifolia	16	5								Brassibase; Rollins 1966
Physaria arenosa	20	5		10, 18						Brassibase; Rollins 1973
Physaria argyrea subsp argyrea	23	18	6,7,8,9,12,15,16,17,18	18,30						Salywon *et al* 2022; Rollins 1966
Physaria argyrea subsp diffusa	27	7,8		14, 16						Warwick *et al* 2006, CCDB
Physaria arizonica	32	5	10	10						Rollins 1966; Brassibase: Rollins 1971
Physaria aurea	34	7								Ward 1983; Brassibase
Physaria bellii	35	4		8						Salywon *et al* 2022, Lysak *et al*., 2009; Brassibase
Physaria brassicoides	38	8		16						Rollins 1993; Brassibase
Physaria calcicola	39	8		20						Rollins & Shaw (1973), Brassibase
Physaria carinata subsp. carinata	43					*No data available*				
Physaria chambersii	46		4,5,8,12	8,16,24	8,10		16		24	Salywon *et al*., 2022, Brassibase; Rollins 1966
Physaria condensata	49									
Physaria congesta	50									
Physaria crassistigma	51					*No data available*				
Physaria curvipes	52									

Physaria densiflora	56		7	14				Warwick 2006 referred to Rollins 1966, CCDB
Physaria densiflora subsp densiflora	57				*No data available*			
Physaria didymocarpa subsp didymocarpa	59			16,24,56				FNA
Physaria didymocarpa subsp lanata	60			8,16				Warwick 2006 referers Mulligan;CCDB
Physaria didymocarpa subsp lyrata	61			24				Warwick 2006 referrers Mulligan; CCDB
Physaria dornii	62				*No data available*			
Physaria douglassii	63		5,15	10,30				FNA, Brassibase; Rollins 1993, Rollins and Shaw, 1973
Physaria eburniflora	64				*No data available*			
Physaria engelmannii	66		6,12,18	12,24,36				FNA, Brassibase; Rollins 1966, Rollins 1993
Physaria eriocarpa	67				*No data available*			
Physaria fendleri	71	6		12	12	24		Salywon *et al*., 2022
Physaria floribunda	72			8				Salywon *et al*., 2022
Physaria floribunda subsp floribunda	75	4		8	8,10	24	16	Salywon *et al*., 2022, FNA
Physaria floribunda subsp osterhoutii	77				*No data available*			
Physaria fremontii	78							
Physaria geyeri subsp purpurea	79			8				Rollins, 1971
Physaria goodingii	80				*No data available*			
Physaria gordonii	83	6		12,32				Salywon *et al.*, 2022
Physaria gracilis subsp gracilis	84			12				FNA
Physaria gracilis subsp nuttallii	85	6		12				Salywon *et al*., 2022
Physaria grahamii	88							
Physaria hitchcockii subsp confluens (not hitch.)	89							
Physaria hitchcockii subsp confluens	91				*No data available*			
Physaria hitchcockii subsp rubicundula	95							

Physaria hitchcockii subsp tumulosa	96						
Physaria hitchcokii subsp hitchcockii	98						
Physaria humilis	99						
Physaria inflata	100	9		18			Salywon *et al*., 2022
Physaria integrifolia	102			16			FNA
Physaria intermedia	106		18	18,20, 36			FNA: Rollins, 1971
Physaria iveyana	107						
Physaria kaibabensis	109			*No data available*			
Physaria kingii subsp cobrensis	110						
Physaria kingii subsp diversifolia	112	5		10			FNA
Physaria kingii subsp kingii	113			*No data available*			
Physaria kingii subsp latifolia	116	5		20	10	20	Salywon *et al*., 2022
Physaria klausii	117			*No data available*			
Physaria lateralis	118						
Physaria lepidota	119	8		16			Salywon *et al*., 2022: CCDB; Warwick, 2006
Physaria lepidota (= P. chambersii membranacea)	120			*No data available*			
Physaria lesicii	121						
Physaria ludoviciana	122	5,10,15		10,20,30			FNA: Brassibase; Rollins 1993, Rollins 1996
Physaria mcvaughiana	125			12			FNA
Physaria medicinae	126			*No data available*			
Physaria mirandiana	129						
Physaria montana	130	5		10			Salywon *et al* 2022
Physaria navajoensis	133			*No data available*			
Physaria nelsonii	136						
Physaria newberryi subsp newberryi	140	4	4,5	8		16	Salywon *et al* 2022

Physaria newberryi subsp yesicola	144	*No data available*			
Physaria obcordata	145				
Physaria occidentalis subsp occidentalis	146		10		FNA
Physaria occidentalis subsp ?????	147	*No data available*			
Physaria ovalifolia subsp ovalifolia	154		12, 24, 36, 48, 50, 72.		FNA
Physaria pachyphylla	157	*No data available*			
Physaria parviflora	158				
Physaria parvula	161	5,10			Brassibase; Rollins 1973, Rollins 1993
Physaria pinetorum	164	5	10		FNA
Physaria prostrata	165	*No data available*			
Physaria prostrata??	167				
Physaria pruinosa	168				
Physaria carinata subsp. pulchella	171				
Physaria pulvinata	175				
Physaria recurvata	186	5	10		Salywon *et al*., 2022: FNA
Physaria reediana	187	5	10, 12		Brassibase; Rollins 1993, Rollins 1996
Physaria rollinsii	190		8		FNA
Physaria saximontana subsp dentata	191	*No data available*			
Physaria saximontana subsp saximontana	192				
Physaria schnaffneri	195	6			Brassibase; Rollins and Rudenburg 1977
Physaria scrotiformis	197	*No data available*			
Physaria sessilis	198		12		FNA
Physaria spathulata	201	*No data available*			
Physaria subumbellata	202		10		FNA
Physaria valida??	207		10		FNA

Physaria vicina	208	*No data available*				
Physaria vitulifera	212	4	8	8	16	Salywon *et al*., 2022
Physaria medicinae (says vitulifera)	213	*No data available*				
Physaria wardii	215	6	12			Rollins and Shaw 1973; Rollins, 1966
Physaria globosa		7	14	14		Salywon *et al.,* 2022
Physaria filiformis			14			Brassibase
Physaria gracilis		6	12			Brassibase; Lysak *et al* 2009, Rollins 1966
Physaria ouachitensis		*No data available*				

Appendix 2.1: Collection information for the samples of *Physaria* species included in the study.

Species	Sample number	Collection Number	Collector	State (or Country)	County (or State)	Elevation (ft.)
Paysonia grandiflora	5	7518	O'Kane & Grady	Texas	Burnet	936
Paysonia lasiocarpa subsp berlandieri	6	7535	O'Kane & Grady	Mexico	Nuevo Leon	4298
Physaria "cnema"	7	8870	O'Kane	Colorado	Dolores	7785
Physaria acutifolia	8	48	Ratcliff & O'Kane	Wyoming	Carbon	4716
Physaria acutifolia	9	4510	O'Kane	Montana	Carbon	5000
Physaria acutifolia	10	8917	Reveal & Broome	Colorado	Dolores	
Physaria acutifolia	11	9061	O'Kane, Lees & Stuart	Utah	San Juan	6442
Physaria alpestris	12	9861	O'Kane & Minnaert-Grote	Washington	Kittitas	3340
Physaria alpestris	13	9862	O'Kane & Minnaert-Grote	Washington	Kittitas	3425
Physaria alpina	14	3982	O'Kane	Colorado	Park	12500
Physaria alpina	15	10156	O'Kane	Colorado	Gunnison	12200
Physaria angustifolia	16	7506	O'Kane & Grady	Oklahoma	McCurtain	496
Physaria angustifolia	17	7507	O'Kane & Grady	Oklahoma	Choctaw	461
Physaria angustifolia	18	7506(D)	O'Kane & Grady	Oklahoma	McCurtain	496
Physaria arenosa	19	3784	O'Kane	Wyoming	Uinta	7000
Physaria arenosa	20	9867	O'Kane & Ratcliff	Wyoming	Albany	7530
Physaria argyrea subsp argyrea	21	7521	O'Kane & Grady	Texas	Gillespie	1363
Physaria argyrea subsp argyrea	22	7531	O'Kane & Grady	Mexico	Nuevo Leon	1308
Physaria argyrea subsp argyrea	23	7528	O'Kane & Grady	Texas	Dimmit	694
Physaria argyrea subsp diffusa	24	7534	O'Kane & Grady	Mexico	Nuevo Leon	5076
Physaria argyrea subsp diffusa	25	7536	O'Kane & Grady	Mexico	Nuevo Leon	7070

Physaria argyrea subsp diffusa	26	7542	O'Kane & Grady	Mexico	Cohuila	6252
Physaria argyrea subsp diffusa	27	7548	O'Kane & Grady	Mexico	Zacatecas	7295
Physaria arizonica	28	160	Grady	Arizona	Coconino	5741
Physaria arizonica	29	169	Grady	Arizona	Yavapai	4603
Physaria arizonica	30	4208	O'Kane & Windham	Arizona	Mohave	4675
Physaria arizonica	31	9057	O'Kane, Lees & Stuart	Arizona	Coconino	7041
Physaria arizonica	32	9059	O'Kane, Lees & Stuart	Arizona	Coconino	6930
Physaria aurea	33	3824	O'Kane	New Mexico	Otero	8330
Physaria aurea	34	9582	O'Kane & Heil	New Mexico	Otero	9130
Physaria bellii	35	3755	O'Kane	Colorado	Boulder	5670
Physaria brassicoides	36	53	Ratcliff & O'Kane	Wyoming	Crook	3915
Physaria brassicoides	37	54	Ratcliff & O'Kane	Wyoming	Crok	4260
Physaria brassicoides	38	7902	O'Kane & Grady	South Dakota	Pennington	3123
Physaria calcicola	39	3753	O'Kane	Colorado	Fremont	6100
Physaria calcicola	40	4487	O'Kane	Colorado	Fremont	5500
Physaria calcicola	41	4488	O'Kane	Colorado	El Paso	6300
Physaria carinata subsp. carinata	42	3796	O'Kane	Montana	Granite	4300
Physaria carinata subsp. carinata	43	3797	O'Kane	Montana	Granite	5600
Physaria chambersii	44	4192	O'Kane & Windham	Utah	Kane	5275
Physaria chambersii	45	4203	O'Kane & Windham	Arizona	Mohave	5000
Physaria chambersii	46	9835	O'Kane & Minnaert-Grote	Nevada	Nye	7200
Physaria condensata	47	40	Ratcliff & O'Kane	Wyoming	Lincoln	7299
Physaria condensata	48	41	Ratcliff & O'Kane	Wyoming	Lincoln	7241
Physaria condensata	49	3783	O'Kane	Wyoming	Sweetwater	7000
Physaria congesta	50	3765	O'Kane	Colorado	Rio Blanco	6050

Physaria crassistigma	51	23145 (BAA)	Ruiz Lea	Argentina	Mendoza	
Physaria curvipes	52	7903	O'Kane & Grady	Wyoming	Sheridan	5348
Physaria curvipes	53	7905	O'Kane & Grady	Wyoming	Sheridan	9193
Physaria curvipes	54	7917 (7919?)	O'Kane & Grady	Wyoming	Big Horn	9500
Physaria densiflora	55	7517	O'Kane & Grady	Texas	Burnet	1092
Physaria densiflora	56	7519	O'Kane & Grady	Texas	Burnet	936
Physaria densiflora subsp densiflora	57	7554	O'Kane & Grady	Mexico	Coahuila	2642
Physaria didymocarpa subsp didymocarpa	58	3794	O'Kane	Montana	Granite	4100
Physaria didymocarpa subsp didymocarpa	59	47	Ratcliff & O'Kane	Wyoming	Gallatin	5013
Physaria didymocarpa subsp lanata	60	52	Ratcliff & O'Kane	Wyoming	Johnson	6343
Physaria didymocarpa subsp lyrata	61	45	Ratcliff & O'Kane	Idaho	Lemhi	4535
Physaria dornii	62	42	Ratcliff & O'Kane	Wyoming	Lincoln	6890
Physaria douglassii	63	4496	O'Kane	Washington	Wenatchee	800
Physaria eburniflora	64	37	Ratcliff & O'Kane	Wyoming	Fremont	8450
Physaria eburniflora	65	38	Ratcliff & O'Kane	Wyoming	Natrona	6075
Physaria engelmannii	66	7514	O'Kane & Grady	Texas	Lampasas	1232
Physaria eriocarpa	67	7927	O'Kane & Grady	Montana	Beaverhead	8700
Physaria fendleri	68	4529	O'Kane	Utah	San Juan	4700
Physaria fendleri	69	7539	O'Kane & Grady	Mexico	Coahuila	8090
Physaria fendleri	70	7558	O'Kane & Grady	Mexico	Coahuila	5327
Physaria fendleri	71	7563	O'Kane & Heil	New Mexico	San Juan	5942
Physaria floribunda	72	10158	O'Kane	Colorado	Saguache	9675
Physaria floribunda subsp floribunda	73	3736	O'Kane & Anderson	Colorado	Mineral	9000

Physaria floribunda subsp floribunda	74	3751	O'Kane	Colorado	Mineral	8800
Physaria floribunda subsp floribunda	75	8249	O'Kane & Heil	New Mexico	Taos	8300
Physaria floribunda subsp osterhoutii	76	3756	O'Kane	Colorado	Summitt	8000
Physaria floribunda subsp osterhoutii	77	3759	O'Kane	Colorado	Grand	8000
Physaria fremontii	78	4495	O'Kane	Wyoming	Fremont	8200
Physaria geyeri subsp purpurea	79	3792	O'Kane	Idaho	Custer	6000
Physaria goodingii	80	9469(B)	O'Kane & Heil	New Mexico	Catron	6060
Physaria gordonii	81	4520	O'Kane & Hutchinson	Oklahoma	Woods	1500
Physaria gordonii	82	3830	O'Kane	New Mexico	Lincoln	5900
Physaria gordonii	83	86	Grady & O'Kane???			
Physaria gracilis subsp gracilis	84	7512B	O'Kane & Grady	Texas	Cooke	881
Physaria gracilis subsp nuttallii	85	7508(A)	O'Kane & Grady	Oklahoma	Marshall	753
Physaria grahamii	86	9832	O'Kane & Minnaert-Grote	Utah	Utah	6600
Physaria grahamii	87	29637	Goodrich	Utah	Duchesne	7567
Physaria grahamii	88	9833	O'Kane & Minnaert-Grote	Utah	Utah	7390
Physaria hitchcockii subsp confluens (not hitch.)	89	00-161	Windham	Nevada	Nevada	8750
Physaria hitchcockii subsp confluens (not hitch.)	90	00-180	Windham	Nevada	Nevada	10250
Physaria hitchcockii subsp confluens (not hitch.)	91	s.n. (98-231)	Windham	Nevada	Nevada	7750
Physaria hitchcockii subsp rubicundula	92	4187	O'Kane & Windham	Utah	Garfield	7500
Physaria hitchcockii subsp rubicundula	93	4719	O'Kane	Utah	Kane	7720
Physaria hitchcockii subsp rubicundula	94	4714	O'Kane	Utah	Garfield	8020

Physaria hitchcockii subsp rubicundula	95	4720	O'Kane	Utah	Iron	8390
Physaria hitchcockii subsp tumulosa	96	4191	O'Kane & Windham	Utah	Kane	5200
Physaria hitchcockii subsp tumulosa	97	4193(A)	O'Kane & Windham	Utah	Kane	5750
Physaria hitchcokii subsp hitchcockii	98	4715	O'Kane	Utah	Garfield	10200
Physaria humilis	99	3800	O'Kane	Montana	Ravalli	8675
Physaria inflata	100	7540	O'Kane & Grady	Mexico	Nuevo Leon	3247
Physaria integrifolia	101	43	Ratcliff & O'Kane	Wyoming	Lincoln	6461
Physaria integrifolia	102	44	Ratcliff & O'Kane	Wyoming	Lincoln	5744
Physaria intermedia	103	4718	O'Kane	Utah	Garfield	7050
Physaria intermedia	104	5504	O'Kane	Utah	Navajo	8000
Physaria intermedia	105	8628	O'Kane & Heil	New Mexico	Sante Fe	6896
Physaria intermedia	106	9834	O'Kane & Minnaert-Grote	Utah	Garfield	8090
Physaria iveyana	107	9056	O'Kane & Heil	New Mexico	Bernallio	10690
Physaria kaibabensis	108	5852	O'Kane	Arizona	Coconino	8606
Physaria kaibabensis	109	4213	O'Kane	Arizona	Coconino	8300
Physaria kingii subsp cobrensis	110	92	Minnaert-Grote & O'Kane	Nevada	Humboldt	5847
Physaria kingii subsp diversifolia	111	93	Minnaert-Grote & O'Kane	Oregon	Wallawa	4870
Physaria kingii subsp diversifolia	112	94	Minnaert-Grote & O'Kane	Oregon	Wallawa	5090
Physaria kingii subsp kingii	113	91	Minnaert-Grote & O'Kane	Nevada	Nye	7310
Physaria kingii subsp latifolia	114	84	Minnaert-Grote & O'Kane	Arizona	Coconino	7620
Physaria kingii subsp latifolia	115	86(b)	Minnaert-Grote & O'Kane	Utah	Iron	10359
Physaria kingii subsp latifolia	116	89	Minnaert-Grote & O'Kane	Nevada	Lincoln	6436
Physaria klausii	117	3799	O'Kane	Montana	Lewis and Clark	5660

Physaria lateralis	118	160 (BAA)	E O. Roig 160	Argentina	Mendoza	
Physaria lepidota	119	4709	O'Kane	Utah	Kane	7130
Physaria lepidota (= P. chambersii membranacea)	120	4722	O'Kane	Utah	Garfield	7590
Physaria lesicii	121	4509	O'Kane	Montana	Carbon	8000
Physaria ludoviciana	122	46	Grady & O'Kane	Wyoming	Fremont	7773
Physaria ludoviciana	123	3772	O'Kane	Utah	Uintah	5400
Physaria ludoviciana	124	7016	O'Kane	Arizona	Coconino	5625
Physaria mcvaughiana	125	7562	O'Kane & Grady	Texas	Pecos	4572
Physaria medicinae	126	10151	O'Kane	Wyoming	Carbon	8475
Physaria mirandiana	127	7537	O'Kane & Grady	Mexico	Nuevo Leon	7045
Physaria mirandiana	128	7551	O'Kane & Grady	Mexico	San Luis Potosi	7138
Physaria mirandiana	129	7538	O'Kane & Grady	Mexico	Coahuila	6616
Physaria montana	130	8630	O'Kane & Heil	New Mexico	Taos	6890
Physaria montana	131	8798	O'Kane, Heil & Mietty	New Mexico	Taos	7080
Physaria montana	132	9866	O'Kane & Minnaert-Grote	Wyoming	Albany	8180
Physaria navajoensis	133	4710	O'Kane	Utah	Kane	7200
Physaria navajoensis	134	5460	O'Kane	New Mexico	McKinley	7475
Physaria navajoensis	135	4220B	O'Kane	New Mexico	McKinley	7400
Physaria nelsonii	136	49	Grady & O'Kane	Wyoming	Sublette	7306
Physaria nelsonii	137	75	Grady & O'Kane	Wyoming	Platte	4957
Physaria nelsonii	138	4493	O'Kane	Wyoming	Sweetwater	7600
Physaria nelsonii	139	54B	Grady & O'Kane	Idaho	Custer	5610
Physaria newberryi subsp newberryi	140	4231	O'Kane & Roth	New Mexico	McKinley	7400
Physaria newberryi subsp newberryi	141	5651	O'Kane	New Mexico	San Juan	6070
Physaria newberryi subsp newberryi	142	8679	O'Kane	New Mexico	McKinley	6880

Physaria newberryi subsp newberryi [racemosa]	143	116	Lees & O'Kane			
Physaria newberryi subsp yesicola	144	34211	Heil & O'Kane	New Mexico	Cibola	6960
Physaria obcordata	145	3762	O'Kane	Colorado	Rio Blanco	6100
Physaria occidentalis subsp occidentalis	146	95	Minnaert-Grote & O'Kane	Idaho	Butte	7918
Physaria occidentalis subsp ?????	147	4503	O'Kane	Idaho	Custer	8800
Physaria occidentalis subsp occidentalis	148	4498	O'Kane	Oregon	Grant	3900
Physaria occidentalis subsp occidentalis	149	4501	O'Kane	Idaho	Blaine	7200
Physaria occidentalis subsp occidentalis	150	4502	O'Kane	Idaho	Custer	8934
Physaria ovalifolia subsp alba	151	7510	O'Kane & Grady	Oklahoma	Carter	1219
Physaria ovalifolia subsp ovalifolia	152	4522	O'Kane	New Mexico	Union	5200
Physaria ovalifolia subsp ovalifolia	153	8758	O'Kane & Heil	New Mexico	Mora	5180
Physaria ovalifolia subsp ovalifolia	154	9053	O'Kane & Heil	New Mexico	Quay	4290
Physaria pachyphylla	155	4511	O'Kane	Montana	Carbon	4760
Physaria pachyphylla	156	7921	O'Kane & Grady	Montana	Carbon	5069
Physaria pachyphylla	157	7925	O'Kane & Grady	Montana	Carbon	5031
Physaria parviflora	158	3766	O'Kane	Colorado	Rio Blanco	6300
Physaria parvula	159	44(X)	Grady & O'Kane	Utah	Daggett	6684
Physaria parvula	160	3782	O'Kane	Utah	Dagget	6750
Physaria parvula	161	4551	O'Kane	Colorado	Grand	9000
Physaria pinetorum	162	3825	O'Kane	New Mexico	Otero	8200
Physaria pinetorum	163	9097	O'Kane & Heil	New Mexico	Torrence	9284
Physaria pinetorum	164	9239	O'Kane & Heil	New Mexico	Luna	5310

Physaria prostrata	165	3790	O'Kane	Utah	Rich	6600
Physaria prostrata	166	4500	O'Kane	Idaho	Custer	6500
Physaria prostrata??	167	3791	O'Kane	Idaho	Custer	6450
Physaria pruinosa	168	3738	O'Kane & Anderson	Colorado	Archuleta	7100
Physaria pruinosa	169	4156	O'Kane	Colorado	Archuleta	7150
Physaria pruinosa	170	4822B	O'Kane & Heil	New Mexico	Rio Arriba	7585
Physaria carinata subsp. pulchella	171	339	Grady	Idaho	Lemhi	6965
Physaria carinata subsp. Pulchella	172	7948	O'Kane & Grady	Montana	Beaverhead	6600
Physaria carinata subsp. pulchella	173	7951	O'Kane & Grady	Idaho	Lemhi	9595
Physaria carinata subsp. pulchella	174	8528	Reveal & Broome	Colorado		
Physaria pulvinata	175	8530	Reveal & Broome	Colorado	Dolores	7685
Physaria pulvinata	176	8532	Reveal & Broome (Isotype)	Colorado	San Miguel	7740
Physaria pulvinata	177	8871	O'Kane, Reveal & Reveal	Colorado	Dolores	7651
Physaria purpurea	178	3821	O'Kane	New Mexico	Otero	5000
Physaria purpurea	179	9058	O'Kane, Lees & Stuart	Arizona	Coconino	6930
Physaria purpurea	180	9134	O'Kane & Heil	New Mexico	Doña Ana	5020
Physaria purpurea	181	9765	O'Kane & Heil	New Mexico	Chaves	5080
Physaria rectipes	182	110	Grady	New Mexico	San Juan	7451
Physaria rectipes	183	117	Grady	Utah	San Juan	6854
Physaria rectipes	184	3802	O'Kane	New Mexico	Rio Arriba	7300
Physaria recurvata	185	7522	O'Kane & Grady	Texas	Gillespie	1880
Physaria recurvata	186	7524	O'Kane & Grady	Texas	Gillespie	2066
Physaria reediana	187	72	Grady & O'Kane	Wyoming	Crook	4062
Physaria reediana	188	98	??Grady & O'Kane			
Physaria reediana	189	7953	O'Kane & Grady	Nebraska	Kimball	7116
Physaria rollinsii	190	3731	O'Kane & Anderson	Colorado	Gunnison	5825

Physaria saximontana subsp dentata	191	56	Ratcliff & Horsch	Montana	Glacier	7193
Physaria saximontana subsp saximontana	192	39	Ratcliff & O'Kane	Wyoming	Fremont	6819
Physaria saximontana subsp saximontana	193	50(A)	Ratcliff & O'Kane	Wyoming	Hot Springs	6709
Physaria saximontana subsp saximontana	194	50(B)	Ratcliff & O'Kane	Wyoming	Hot Springs	6709
Physaria schnaffneri	195	7546	O'Kane & Grady	Mexico	Durango	6131
Physaria schnaffneri	196	7547	O'Kane & Grady	Mexico	Zacatecas	7044
Physaria scrotiformis	197	7977	O'Kane, Heil & Jamieson	Colorado	LaPlata	11827
Physaria sessilis	198	7523	O'Kane & Grady	Texas	Gillespie	2066
Physaria spathulata	199	49	Ratcliff & O'Kane	Wyoming	Carbon	4716
Physaria spathulata	200	7923	O'Kane & Grady	Montana	Carbon	8712
Physaria spathulata	201	9872	O'Kane & Ratcliff	Wyoming	Crook	4125
Physaria subumbellata	202	73	Grady & O'Kane	Wyoming	Crook	5296
Physaria subumbellata	203	131	Grady	Utah	Uintah	7759
Physaria subumbellata	204	3777A	O'Kane	Utah	Uintah	7550
Physaria valida??	205	3827	O'Kane	New Mexico	Lincoln	7130
Physaria valida??	206	3842	O'Kane	New Mexico	Eddy	6300
Physaria valida??	207	3845	O'Kane	New Mexico	Eddy	6300
Physaria vicina	208	3722	O'Kane & Anderson	Colorado	Ouray	6900
Physaria vitulifera	209	3701	O'Kane	Colorado	Jefferson	5070
Physaria vitulifera	210	3752	O'Kane	Colorado	Fremont	6400
Physaria vitulifera	211	4925	O'Kane	Colorado	Costilla	8275
Physaria vitulifera	212	10159	O'Kane	Colorado	Jefferson	6950
Physaria medicinae (says vitulifera)	213	34	Ratcliff & O'Kane	Wyoming	Carbon	8083
Physaria vitulifera/floribunda?	214	9196	O'Kane & O'Kane	Colorado	Chaffee	9640
Physaria wardii	215	4211	O'Kane	Arizona	Coconino	7050

Physaria wardii	216	4212	O'Kane	Arizona	Coconino	7000
Physaria wardii	217	4721	O'Kane	Utah	Iron	10420
Physaria wyndii	218	7557	O'Kane & Grady	Mexico	Coahuila	5327
Physaria globosa	219	360	Edwards, Linan and Dell	Tennessee		
Physaria globosa	220	362	Edwards, Linan and Dell	Tennessee		
Physaria gracilis	221	Tx1 -15		Texas		
Physaria gracilis	222	Tx1 -20		Texas		
Physaria gracilis	223	Tx19		Texas		
Physaria ouachatensis	224	Ar1.19		Arkansas		
Physaria ouachatensis	225	Ar 312		Arkansas		
Physaria ouachatensis	226	Ar 009		Arkansas		
Physaria ouachatensis	227	Ar 2.17		Arkansas		
Physaria filiformis	228	MO 013		Missouri		
Physaria filiformis	229	MO 4.15		Missouri		

Supplementary information S1: Parameters used in the ipyrad assembly for the final dataset.

Parameters	Description
denovo	[5] [assembly_method]: Assembly method (denovo, reference)
2brad	[7] [datatype]: Datatype (see docs): rad, gbs, ddrad, etc.
4	[9] [max_low_qual_bases]: Max low quality base calls (Q<20) in a read
33	[10] [phred_Qscore_offset]: phred Q score offset (33 is default and very standard)
10	[11] [mindepth_statistical]: Min depth for statistical base calling
4	[12] [mindepth_majrule]: Min depth for majority-rule base calling
8000	[13] [maxdepth]: Max cluster depth within samples
0.9	[14] [clust_threshold]: Clustering threshold for de novo assembly
0	[15] [max_barcode_mismatch]: Max number of allowable mismatches in barcodes
0	[16] [filter_adapters]: Filter for adapters/primers (1 or 2=stricter)
35	[17] [filter_min_trim_len]: Min length of reads after adapter trim
2	[18] [max_alleles_consens]: Max alleles per site in consensus sequences
0.02	[19] [max_Ns_consens]: Max N's (uncalled bases) in consensus
0.05	[20] [max_Hs_consens]: Max Hs (heterozygotes) in consensus
5	**[21] [min_samples_locus]: Min # samples per locus for output**
0.083	[22] [max_SNPs_locus]: Max # SNPs per locus
1	[23] [max_Indels_locus]: Max # of indels per locus
0.5	[24] [max_shared_Hs_locus]: Max # heterozygous sites per locus
0, 0, 0, 0	[25] [trim_reads]: Trim raw read edges (R1>, <R1, R2>, <R2) (see docs)
0, 0, 0, 0	[26] [trim_loci]: Trim locus edges (see docs) (R1>, <R1, R2>, <R2)
l, p, s, n, k, a, g, G, u, v, t, m	[27] [output_formats]: Output formats (see docs)